

XX 25-SEP-2003.
 PD 17-MAR-2003; 2003WO-EP002735.
 XX 20-MAR-2002; 2002DE-01012892.
 PR (BADI) BASF PLANT SCI GMBH.
 XX Kock M, Bauer J;
 PI WPI: 2003-803889/75.
 XX N-PSDB; ADH89337.
 DR Reducing expression of at least two target genes, useful e.g. for
 PT producing transgenic plants, using partly double-stranded interfering
 PT RNA.
 PS Disclosure; SEQ ID NO 113; 228bp; German.
 XX This invention describes a novel method for reducing the expression of at
 CC least two different endogenous target genes in a eukaryotic cell or
 CC organism by introducing an RNA molecule that is at least partly double
 CC stranded. The transcribed RNAs from at least two target genes have
 CC homology below 90% and the RNA molecule is formed as a single, self-
 CC complementary molecule. At least one of the double-stranded structures
 CC formed from individual sense sequences has an even number of repeats of
 CC 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At
 CC least two target genes are selected from different classes of storage
 CC protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-
 CC prolamine and at least one of the sense sequences is identical to storage
 CC protein sequences or genes in the homogenistate metabolic pathway or
 CC enzyme types, e.g. acetyl transacylases, cholesteraes, (de)branching
 CC enzymes or cellulases. The RNA of the invention, also related cassettes,
 CC expression systems, vectors and transgenic organisms are used for
 CC preparation of pharmaceuticals, in biotechnological processes and plant
 CC biotechnology, specifically in plants to improve protection against
 CC abiotic stress, to modify composition and/or content of fatty acids,
 CC lipids and oils, to modify carbohydrate composition, to alter colour or
 CC pigmentation, to reduce content of storage proteins, to increase
 CC resistance to pathogens, to inhibit stem break, to delay fruit ripening
 CC or aging, to induce male sterility, to reduce content of toxic or
 CC unwanted components, to modify lignification and/or lignin content, to
 CC modify the fibre component in foods or fibre quality in cotton, to reduce
 CC susceptibility to shock, to increase synthesis of Vitamin E, to reduce
 CC contents of nicotine, caffeine or theophylline and to increase methionine
 CC content, by reducing chreonine biosynthesis. The method provides a rapid
 CC and efficient way of reducing gene expression, can inhibit more than one
 CC target gene, prevents development of multiple phenotypes (since the
 CC transcription rate is the same for all RNA sequences, significantly
 CC reducing the selection process required to produce an organism with
 CC effective suppression of all target genes), avoids problems of epigenetic
 CC gene silencing, does not require synthesis of individual RNA sequences,
 CC and the method can be applied to plants with complex (polyploid) genomes.
 CC No interference between the individual RNA sequences occur. This sequence
 CC represents a protein encoded by a target gene used in the method of the
 CC invention.
 XX
 SQ Sequence 307 AA;
 Query Match 30.8%; Score 440.5; DB 7; Length 307;
 Best Local Similarity 42.8%; Pred. No. 1,4e-29;
 Matches 125; Conservative 37; Mismatches 75; Indels 55; Gaps 16;
 QY 13 PSQGGPQEQVPLVQGGQFP-----GQQGQFPFPQGPYFPQPFPSQGPYLTQGPFPQPRUP 67
 DB 27 PGLERFWQGGQPLPFPQGFPPQGPPLFSQQGQ---QGLFRFGQGSFSGQQGP-----PFWGQGP 78
 QY 68 YPQPGSPFPQGP-----YPQPGYSGPQGP---SQGQAQQGQQGQQGQQGQQGQ--- 114
 DB 79 FSGGQGPPLFPQGPFSQGGQGLVLPQ---QPFPSQGGQGPVLPQSPFPQGGQGHQGLVQGGQIP 137
 QY 115 -LQGLLQGGQILPQMDVYLQGH--NIAH-----ASGQVLAQSTYQLLGLCCGHLWQIPQGS 167

DB 138 VQPSILQQLNFC-KVFLQGGQCSFVAMPQRLARQGLQSSCHVMQGGCCQQLPQIRQGS 196
 QY 168 QCGAIIHVHVAIILHQOQKQGGQSSQVSPFOPLQGGFPLGGSFPSQGNPQ----- 219
 DB 197 RYEAIRAIYVAIL--QEQGQVGSISQGGQGPQ---LGGCVSQPGQGSQQLGGQGPQ 251
 QY 220 ---AGGS-VQPGQPLPQPEIRNALQTLPMCMNYIAPY--CTIAPGIGT 265
 DB 252 QQLAQGTFLQPHQIAQLEVMVTISALRIPLTWCSSVNPVPLRTTTSVPPGV-GT 302
 RESULT 8
 ID ADG44134
 XX ADG44134 standard; protein; 307 AA.
 AC ADG44134;
 XX 26-FEB-2004 (first entry)
 DT T. aestivum glutenin-1D1 protein.
 DE oil content; plant; storage protein; seed-specific promoter; 2S-albumin;
 XX 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;
 KW oil production; fat production; free fatty acid production; food;
 KW animal feed; pharmaceutical; fine chemical production; glutenin.
 XX
 CS Triticum aestivum.
 XX WO2003077643-A2.
 XX 25-SEP-2003.
 XX 17-MAR-2003; 2003WO-EP002735.
 PF 20-MAR-2002; 2002DE-01012893.
 XX (BADI) BASF PLANT SCI GMBH.
 PA Bauer J;
 XX WPI: 2004-011485/01.
 DR N-PSDB; ADG44133.
 DR Increasing total oil content of plants, useful e.g. as foods or animal
 PT feeds, by reducing amount of storage proteins, particularly with double-
 PT stranded interfering RNA.
 PS Claim 4; SEQ ID NO 174; 253bp; German.
 XX This invention describes a novel method for increasing the total oil
 CC content of a plant by reducing the amount of at least one storage protein
 CC in the plant (or its tissue, organs, parts or cells) and selecting plants
 CC that have higher total oil content than starting plants. The storage
 CC protein is suppressed by introducing antisense RNA, optionally combined
 CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding
 CC factors directed against storage protein genes, viral sequences that
 CC recombination of endogenous storage protein genes or mutations into
 CC storage protein genes. Most preferably a plant cell is stably transfected
 CC with a recombinant expression construct, then regenerated to plants that
 CC express the incorporated sequence. The expression constructs particularly
 CC contain a seed-specific promoter and they are introduced into plants by
 CC standard methods, e.g. via Agrobacterium. The preferred storage proteins
 CC of the invention are 2S-albumens, 7S or 11S/12S-globulins or zein-
 CC prolamines. Transgenic organisms produced by the new method are used for
 CC production of oils, fats, free fatty acids or their derivatives, useful
 CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence
 CC represents a storage protein used to illustrate the method of the
 CC invention.
 XX
 SQ Sequence 307 AA;

Query Match 30.8%; Score 440.5; DB 8; Length 307;
 Best Local Similarity 42.8%; Pred. No. 1.4e-29;
 Matches 125; Conservative 37; Mismatches 75; Indels 55; Gaps 16;

13 PSQQQPCQVPLVQCCQCF-----GQQQPPPPQPPYQPPQPPSQQPPYLQPPPPRLP 67
 27 PGLERPWQQQPLPQQTTPQQPLPSQQQ---QQLFPQPPSQQQP-----PFWQQQPP 78
 68 YPQPSPPPPQPP-----YPQPPQPSQPPPI---SQQQAQQQQQQQQQQQQQQQI- 114
 79 FSQQQPLLPPQPPSPSQQQQLVLPQ-QPPFSQQQPPVPPPPQSPPPQQQGHQQLVQQQLP 137
 115 -LQQLLQQLLPCMDVVLQOH--NIH---ARSQVLAQSTYQLQELCCGHLWLPQSQS 167
 138 VQPSILQQLMPC-KVFLQCCQSPYPMQRLARSQMLQSSCHVWQQCCQQLPQIPQGS 196
 168 QCAIHNVNVAIILHQCKQCCQSSQVSPFQPP-LQVPLCGGSRPSQPPQ----- 219
 197 RYEA-RAIIVSIIL--QCCQVQGSIDQCCQPPQ---LQCCVSPQPPQSSQQLGQPPQ 251
 220 ---AGGS-VQPPQPLPQPEIRNLALQTLPMACNYIAPY--CTIAPFGIFGT 265
 252 QQLAGTFLQPHQIAQLEVMISIALRLIPLMCSVNVPLKRTTTSVPPGV-CT 302

RESULT 9

AD071661
 AD071661 standard; protein; 298 AA.

AD071661;

12-AUG-2004 (first entry)

Amino acid sequence of glutenin clone LMW6.

low molecular weight subunit; LMW subunit; glutenin;

wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;

gluten intolerance.

Triticum sp.

Location/Qualifiers

Disulfide-bond 25 /note = this residue forms an intermolecular disulfide

Disulfide-bond 227 /note = this residue forms an intermolecular disulfide

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

Disulfide-bond 227

The present sequence represents a low molecular weight (LMW) subunit of
 CC glutenin, and is derived from clone LMW6 which is isolated from wheat
 CC cultivar Cheyenne. The LMW6 subunit does not contain the allergenic
 CC epitope GQQP and shows some minor differences to published sequences.
 CC It therefore represents a new allele for LMW subunit genes. The LMW6
 CC polypeptide was modified to produce modified glutenin polypeptides of the
 CC invention. In these modified polypeptides one or more cysteine residues
 CC responsible for intermolecular cross linking through disulfide bridges
 CC are deleted or substituted. The modified glutenin polypeptide is useful
 CC as a gliadin substitute. It is also useful in the preparation of
 CC foodstuffs, such as flour or for the preparation of pharmaceutical
 CC products, such as tablets, where the foodstuffs contain a considerably
 CC reduced amount of gliadin protein or no gliadin protein. Pharmaceutical
 CC compositions comprising the modified polypeptide of the invention are
 CC useful for treating patients suffering from coeliac disease or persons
 CC who are intolerant to gluten.

Sequence 298 AA;

Query Match 30.6%; Score 437.5; DB 8; Length 298;

Best Local Similarity 41.0%; Pred. No. 2.5e-29;
 Matches 118; Conservative 35; Mismatches 74; Indels 61; Gaps 13;

13 PSQQQPCQVPLVQCCQCF-----GQQQPPPPQPPYQPPQPPSQQPPYLQPPPPRLP 71
 27 PGLERPWQQQPLPQQTTPQQPLPSQQQ---QQLFPQPPSQQQP-----PFWQQQPP 74
 72 GSFPQPPYQPPQPPYSPQPPISQQAQQQQQQQQQQQQQ-----LQQT----- 118
 75 LFSQKQPPVLPQQPAPFSQQQCTVLPQPPAFSQQQGHQQLQQLPQVHSILQQLMPCVVF 134
 115 LQQLLQQLLPCMDVVLQOHNIHARSQVLAQSTYQLQELCCGHLWLPQSQS 178
 135 LQQQ---GSPVAMPQ---LARSQWQSSSCVWQCCQQLPRLPESRYEALTAIIFS 188
 179 IILHQCK-----QCCQSSQVSP-----FQPLQGYPLQGSFPPQ 216
 189 IILHQCKQFPVQPPQCCQVQGVYQPPQSQQQLGQCFQPPQ---LQ---QPPQQ 243
 217 NEQAQGSVPPQQLPQPEIRNLALQTLPMACNYIAPYCTI--APFGI 262
 244 QVQKGTFLQPHQIAQLEVMISIALRLIPLMCSVNVPLYSITSAPLGV 291

RESULT 10

ABB71695
 ABB71695 standard; protein; 1798 AA.

ABB71695;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 41877.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-05009231.

23-MAR-2000; 2000US-0191637P.

11-JUL-2000; 2000US-00614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers BW,

WPI; 2001-656860/75.

DR N-PSDB; ABL15738.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 41877; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1798 AA;
 Query Match 20.2%; Score 288.5; DB 4; Length 1798;
 Best Local Similarity 29.5%; Pred. No. 1.2e-15;
 Matches 102; Conservative 18; Mismatches 91; Indels 135; Gaps 12;
 QY 4 PVVQLQPNPSQQPQEQVPLVQQQPPGQQQF---PPQQPYPPQPPSPQQPYQLQLP 60
 DB 150 PAFQMTPPPPSPQNAQQQHLRQQQWQCNQCHMMGSPQPPQPSQMSFQQQQ--QMQP 207
 QY 61 FPPQ---RLFPQ-----PQSFPQPPY----- 80
 DB 208 FQPPHQQNMQLQQQQQLAQQQQQQQQSPHISFQSPQISQTPPMQAKLHQHVVPQ 267
 QY 81 PQQPQYSQ---PQQPIS-----QQQAQQQQQQQQQ----- 109
 DB 268 ATPQPSQFSQOKPIDPDPVQAVLSRSALSSNODSLIWRQQQLKQQQQQQQQQQMAPQ 327
 QY 110 -QQQQQLLQQLLQQLLPCMMVVLQCHNIAHBSQVLAQSTYQLQLCCQHLMLQEPQSQ 168
 DB 328 PQQQQAQQPQQQQQ---QQQPQQQHTPPSRQSPLOQQQPTPLQLQ-----QQPQQN 375
 QY 169 CQAHHVVAHLLHQQKQKQSPSSQVSPQPPAQYPLGSGSFRPSQ----- 216
 DB 376 AQQI-----QQQQQQQQQQQQQQQQQQQQVLTQQQPQPPQQQVITQBHVINT 425
 QY 217 -----NPPAQGSVPPQQLPQPEIRNL 238
 DB 426 STRAQGGQIIQSHMSLALQKQQLLHVQQAQQQPPQQQQQTTOVQQL 471
 RESULT 11
 ID AAY96255 standard; protein; 1162 AA.
 XX
 AC AAY96255;
 XX
 DT 12-SEP-2003 (revised)
 DT 11-SEP-2000 (first entry)
 XX
 DE Kaposi's sarcoma-associated herpesvirus LANA.
 XX
 KM Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
 KM latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
 KM Human herpes virus 8; HHV8; rhadino virus cis-acting element; RYCAE;
 KM Kaposi's sarcoma; Primary effusion lymphoma; PEL;
 KM human immunodeficiency virus; HIV; multicentric Castlemans disease.
 XX
 OS Human herpesvirus 8.
 XX
 FH Key Location/Qualifiers
 FT Domain 14..17
 FT Domain /note="nuclear localisation signal, NLS"
 FT Domain 64..70

FT FT /note="nuclear localisation signal, NLS"
 FT Region 320..429
 FT /note="acidic repeat region"
 FT Region 430..549
 FT /note="Gln, Glu, Pro-rich region"
 FT Region 550..589
 FT /note="Gln, Glu, Pro, Arg-rich region"
 FT Region 590..759
 FT /note="Gln, Glu, Asp-rich region"
 FT Region 760..840
 FT /note="Gln, Glu-rich region"
 PN WO200029626-A1.
 PD 25-MAY-2000.
 PD 19-NOV-1999; 99WO-US027508.
 PF 19-NOV-1999; 98US-00109422.
 PR 21-APR-1999; 99US-00298568.
 PR 19-NOV-1998; 98US-00109422.
 PR 21-APR-1999; 99US-00298568.
 PA (KIEP/) KIEP E D.
 PA (BALL/) BALLESTAS M E.
 PA (KAYE/) KAYE K M.
 PI Kieff ED, Ballestas ME, Kaye KM;
 XX
 XX WPI; 2000-387829/33.
 DR N-PSDB; AAA30290.
 PT Treating or preventing a disease associated with rhadino virus infection
 PT in a mammal which includes Kaposi's Sarcoma and Primary Effusion
 PT Lymphoma.
 PS Disclosure; Fig 7; 70bp; English.
 PS
 XX
 CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,
 CC (KSHV) latency-associated nuclear antigen (LANA). KSHV is also known as
 CC Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma-2
 CC herpes virus class. The LANA protein is necessary for the efficient
 CC persistence of rhadino virus DNA in mammalian cells. Persistent rhadino
 CC virus infection is implicated in a variety of diseases e.g. Kaposi's
 CC Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
 CC Castlemans disease. In addition, KS is a common malignancy in HIV
 CC patients. KSHV persists in host cells in a latent form. One of the few
 CC genes expressed from the latent viral DNA is LANA. LANA associates with
 CC both human chromosomes and with the rhadino virus cis-acting element
 CC (RYCAE), thereby providing a tethering function: the KSHV DNA episome is
 CC "tied" to the host chromosomes. This allows the viral DNA to persist in
 CC the host cell. The present sequence may be used to screen and identify
 CC molecules that inhibit LANA interaction with RYCAE, thereby interfering
 CC with the latency cycle of this virus. Potential antiviral treatments for
 CC the above mentioned diseases may therefore be based on LANA deregulation.
 CC (Updated on 12-SEP-2003 to standardise OS field)
 CC
 SQ Sequence 1162 AA;
 Query Match 19.1%; Score 273; DB 3; Length 1162;
 Best Local Similarity 36.4%; Pred. No. 1.5e-14;
 Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;

QY 4 PVVQLQPNPSQQPQEQVPLVQ---QQQFPQQQ---QQPFPQQPYPPQPPSPQQPYQLQLP 59
 DB 495 PLOEPQQQCEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEP 553
 QY 60 P---PQQPRLPY---PQ---PQSFPQQPYPPQ---QPPQSPQPPQISQQAQQQQQQQQQQ 109
 DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQ 613
 QY 110 QQQQQLLQQLLQQLLPCMMVVLQCHNIAHBSQVLAQSTYQLQLCCQHLMLQEPQSQ 169
 DB 614 EQQQQEQQQQEQQ-----QDEQQQEQQQQEQQQDEQQQDEQQQEQQEQ-----QQQDEQQQ 663

170 QAHINVVHAILHQQGKQGGSSGVSFGQPPQGYPLGQSGSRFSGQMPQAGQSVDPQQL 220
 Db 664 EQGQD-----EQGQDEQQQDPRQQQDPRQQQDEQQQDDEQQQDDEQQQDEQQQDEQQQD 716
 Oy 230 PQFE 234
 Db 717 EQQD 721
 RESULT 12
 ID AAY58500
 AAAY58500 standard; protein; 1162 AA.
 AC AAY58500;
 DE 06-AUG-2003 (revised)
 DT 10-APR-2000 (first entry)
 DE HHV8 ORF 73 protein, SEQ ID NO:21.
 XX HHV8, detection; diagnosis; Kaposi's sarcoma; AIDS; immunogen; antigen.
 XX Human herpesvirus 8.
 XX Key Location/Qualifiers
 XX Misc-difference 96
 XX /label= unknown
 XX W09961903-A2.
 XX 02-DEC-1999.
 XX 26-MAY-1999; 99MO-US011407.
 XX 26-MAY-1998; 98US-0086695P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Pau C;
 XX WPI; 2000-097142/08.
 New methods and compositions for the detection of human herpesvirus.
 Claim 2; Page 59-62; 68pp; English.
 Sequences AAY58480-Y58532 represent immunogenic polypeptides derived from
 human herpes virus type 8 (HHV8, a gammaherpesvirus). HHV8 plays an
 important role in the pathogenesis of AIDS-related Kaposi's sarcoma. The
 invention relates to a novel method of detecting the presence of human
 herpesvirus 8 in a biological sample using peptides representative of
 dominant antigenic regions of HHV8. The method comprises contacting one
 or more isolated, immunogenic HHV8 peptides with an antibody-containing
 biological sample, and detecting the formation of a complex between the
 peptide and the antibody. The presence of a peptide-antibody complex
 indicates the presence of human herpesvirus 8. The detection of HHV8
 infection can be used to diagnose AIDS-associated Kaposi's sarcoma. The
 HHV8-specific antibodies are useful therapeutically when for the passive
 immunisation of a human against HHV8 infection, thereby reducing HHV8
 related disease. The detection and treatment assays are highly specific, sensitive and
 accurate. Early detection and treatment of Kaposi's sarcoma could
 diminish the severity of symptoms related to AIDS and the sensitive
 techniques could reduce erroneous characterisations of skin disorders.
 Previous assays for HHV8 antibodies such as immunofluorescence assays,
 immunoblots and enzyme immunoassays lack the sensitivity and accuracy
 needed for reliable diagnosis of Kaposi's sarcoma. Further advantages of
 the assays are that reproducible results are obtained and the method is
 suitable for rapid throughput and screening of samples economically.
 (Updated on 06-AUG-2003 to correct OS field.)

	Best Local Similarity 38.4%; Pred. No. 1,56-14;	
	Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10	
QY	4 PVFQLQPNNSQQQPPQSVPLVQ--QQQPFQQ--QQPPQYPPQPPPPSQPYLQLQ 59	
DB	495 PLEPPQDEPPQDEPPQDEPPQDEPPQDEPPQDEPPQDEPPQDEPPQDEPPQDE 553	
QY	60 P--FPQPLPY--PQ--PQSFPPQPYQP-QPYVSPQPPISQQQAQQQQQQQQQQ 109	
DB	554 PQQEPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613	
QY	110 QQQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQLLQQL 169	
DB	614 EQQQDEQQQDEQQQ--QEQQDEQQQDEQQQDEQQQDEQQQDE--QQQDEQQQ 663	
QY	170 QALHNVHAH11HQQQKQQQPPSSGVSTQPLQYPLQGSFPPSQQNPQAQSVQPPQL 229	
DB	664 EQQQD-----EQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQ 716	
QY	230 PPFEE 234	
DB	717 EQQQD 721	
	RESULT 13	
	AAB62331	
ID	AAB62331 standard; protein; 1162 AA.	
XX		
AC	AAB62331;	
XX		
DT	06-AUG-2003 (revised)	
DT	29-JUN-2001 (first entry)	
XX		
DE	Amino acid sequence of KSHV tethering protein LANA.	
XX		
KM	Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;	
KM	Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;	
KM	KSHV; latency-associated nuclear antigen; LANA.	
OS	Human herpesvirus 8.	
XX		
PN	MO200125484-A2.	
XX		
PD	12-APR-2001.	
XX		
PF	29-SEP-2000; 2000MO-US026908.	
XX		
PR	01-OCT-1999; 99US-00410399.	
XX		
PA	(UNMI) UNIV MICHIGAN.	
XX		
PI	Robertson ES, Cotter MA;	
XX		
DR	WPI; 2001-281736/29.	
XX		
PT	N-PSDB; AAF82901.	
XX		
PS		
XX		
CC	Disclosure; Fig 9A; 60bp; English.	
CC	The invention provides a composition comprising nucleic acid, histone H1	
CC	protein and expression vector operatively encoding a protein suitable	
CC	for tethering the nucleic acid to the histone H1 protein, where the	
CC	tethering protein is LANA. The composition is useful in aiding the	
CC	revention of the viral DNA in the host cell. The viral vector encodes a	
CC	protein suitable for tethering DNA to histone H1. Methods for screening	
CC	for compounds which are agonistic or antagonistic for the tethering of	
CC	viral proteins to histone H1 and DNA binding sites are useful for	
CC	developing the method of viral transfer. The composition has applications	
CC	to gene therapy, including the treatment of multiple sclerosis,	
CC	Parkinson's disease, Huntington disease and diabetes. The present	
CC	sequence represents the amino acid sequence of the Kaposi's sarcoma	

CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
 CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
 CC 2003 to correct OS field.)
 CC
 XX
 XX
 SQ Sequence 1162 AA;

Query Match 19.1%; Score 273; DB 4; Length 1162;
 Best Local Similarity 38.4%; Pred. No. 1.5e-14;
 Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;

QY 4 PVPOLQPNPSQQQPPQGVPLVQ--QQQPFQQQ--QQPFPQQPYPPQPPFPSPQSPYQLQ 59
 DB 495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQ 553
 QY 60 P--FPQPLPY---PQ---PQSFPPQPYPP--QPYSQPQQPISQQQAQQQQQQQQQQ 109
 DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQ 613
 QY 110 QQQQILQQILQQILPCMDVYLQGHNTAHARSQVLCQSTYQLQLQELCCGHLMQIPBSQC 169
 DB 614 EQQDDEQQDDEQ---QDEQQDDEQQDDEQQDDEQQDDEQQDDE---QQDDEQQD 663
 QY 170 QAHNVVHAITLHQGQKQQQPSQVSFPQPLQGYPLGQSFPPSQQNFQAGSVQPPQL 229
 DB 664 EQQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
 QY 230 PQQE 234
 DB 717 EQQD 721

RESULT 14

ABB05621
 ID ABB05621 standard; protein; 1162 AA.

XX
 AC ABB05621;

XX
 DT 25-APR-2002 (first entry)

XX
 DE Kaposi's sarcoma-associated herpesvirus LANA protein.

XX
 KW Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAR; PEL;
 KW KSHV terminal repeat; rhadino virus cis acting element; episome;
 KW primary effusion lymphoma; latency-associated nuclear antigen;
 KW gene therapy; gene transfer.

XX
 OS Human herpesvirus 8.

XX
 PN US6322792-B1.

XX
 PD 27-NOV-2001.

XX
 PF 21-APR-1999; 99US-00298568.

XX
 PR 19-NOV-1998; 98US-0109422P.

XX
 PA (KIEF/) KIEFF E D.

XX
 PI Kieff ED, Ballestas ME, Kaye KM;

XX
 DR WPI: 2002-153769/20.

XX
 DR N-PSDB; ABA93487.

XX
 PT System for episomal retention of plasmids in mammalian cells, useful in
 gene therapy, comprises rhadinoviral LANA and RVCAR sequences.

XX
 PS Disclosure; Fig 7; 27pp; English.

XX
 CC The present invention describes a system (A) for maintaining a plasmid as
 CC an episome in mammalian cells, comprising the rhadinoviral sequence LANA
 CC (latency-associated nuclear antigen) of 3469 base pairs (see ABA93487,
 CC S1) expressed in the cell, and the rhadinoviral sequence RVCAR
 CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2).

CC present in the plasmid. Also describes is a method for maintaining a
 CC closed circular DNA in a cell by expressing (S1) in the cells and having
 CC (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
 CC particularly used in gene therapy (or other gene transfer applications)
 CC that uses mammalian cells in which LANA is expressed. (A) improves
 CC persistence of gene therapy vectors in cells. The present sequence
 CC represents Kaposi's sarcoma-associated herpesvirus (KSHV, also called
 CC human herpesvirus 8) LANA protein, which is used in the exemplification
 CC of the present invention
 CC
 XX
 XX
 SQ Sequence 1162 AA;

Query Match 19.1%; Score 273; DB 5; Length 1162;
 Best Local Similarity 38.4%; Pred. No. 1.5e-14;
 Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;

QY 4 PVPOLQPNPSQQQPPQGVPLVQ--QQQPFQQQ--QQPFPQQPYPPQPPFPSPQSPYQLQ 59
 DB 495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQ 553
 QY 60 P--FPQPLPY---PQ---PQSFPPQPYPP--QPYSQPQQPISQQQAQQQQQQQQQQ 109
 DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQ 613
 QY 110 QQQQILQQILQQILPCMDVYLQGHNTAHARSQVLCQSTYQLQLQELCCGHLMQIPBSQC 169
 DB 614 EQQDDEQQDDEQ---QDEQQDDEQQDDEQQDDEQQDDEQQDDE---QQDDEQQD 663
 QY 170 QAHNVVHAITLHQGQKQQQPSQVSFPQPLQGYPLGQSFPPSQQNFQAGSVQPPQL 229
 DB 664 EQQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
 QY 230 PQQE 234
 DB 717 EQQD 721

RESULT 15

ADJ65096
 ID ADJ65096 standard; protein; 1162 AA.

XX
 AC ADJ65096;

XX
 DT 20-MAY-2004 (first entry)

XX
 DE HHV8 latency-associated nuclear antigen, LANA.

XX
 KW HHV8; latency-associated nuclear antigen; LANA;
 KW primary effusion lymphoma; PEL; virucide; gene therapy; KSHV;
 KW Kaposi's sarcoma-associated herpesvirus; episome; RVCAR;
 KW rhadinovirus cis-acting element; rhadinovirus infection.

XX
 OS Human herpesvirus 8.

XX
 PN US2004037847-A1.

XX
 PD 26-FEB-2004.

XX
 PF 28-JUN-2001; 2001US-00894273.

XX
 PR 19-NOV-1998; 98US-0109422P.

XX
 PR 21-APR-1999; 99US-00298568.

XX
 PA (KIEF/) KIEFF E D.

XX
 PA (BALI/) BALLESTAS M E.

XX
 PA (KAYE/) KAYE K M.

XX
 PI Kieff ED, Ballestas ME, Kaye KM;

XX
 DR WPI: 2004-191011/18.

XX
 DR N-PSDB; ADJ65095.

XX
 PT Assays for compounds that modulate rhadino virus LANA action in trans on

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 Seconds
(without alignments)
928,452 Million cell updates/sec

Title: US-10-089-700-3-R65
Perfect score: 1431
Sequence: 1 VAVPVQLOPQNPQQQPE.....CNVYAPYCTIAFCIFGTN 266

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillset.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	464	32.4	369	2	US-08-991-300-2
2	273	19.1	1162	2	US-08-728-323A-2
3	273	19.1	1162	3	US-09-298-568-2
4	273	19.1	1162	4	US-09-410-399-2
5	273	19.1	1162	4	US-09-894-273-2
6	249	17.4	788	2	US-08-918-914-4
7	234	16.4	256	4	US-09-248-796A-21251
8	232	16.2	498	4	US-09-270-767-45042
9	217	15.2	579	4	US-09-668-119-3
10	217	15.2	2074	4	US-09-481-356C-9
11	202	14.1	2023	4	US-09-481-356C-8
12	202	14.1	2124	4	US-09-538-092-1377
13	198.5	13.9	505	4	US-09-248-796A-19253
14	197.5	13.8	663	4	US-09-270-767-41220
15	197.5	13.8	1591	4	US-09-270-767-45698
16	197.5	13.8	2441	1	US-08-194-468-2
17	197.5	13.8	2441	3	US-08-961-739-2
18	197.5	13.8	2441	3	US-09-514-247A-8
19	197.5	13.8	2441	4	US-09-666-316-2
20	196.5	13.7	379	4	US-09-248-796A-23759
21	196.5	13.7	2442	3	US-09-514-247A-10
22	196.5	13.7	2442	4	US-09-538-092-1370
23	196	13.7	216	4	US-09-248-796A-21017
24	195.5	13.7	729	3	US-09-625-188-20
25	188	13.1	295	4	US-09-248-796A-20004
26	186.5	13.0	320	4	US-09-248-796A-24758
27	185	12.9	169	4	US-09-248-796A-28087

28	184.5	12.9	316	4	US-09-270-767-42663	Sequence 42663, A
29	184.5	12.9	332	4	US-09-248-796A-21649	Sequence 21649, A
30	184	12.9	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	1319	4	US-09-538-092-1291	Sequence 1291, Ap
34	177	12.4	261	4	US-09-602-565-34	Sequence 34, Appl
35	176.5	12.3	684	4	US-09-823-240A-9	Sequence 15319, A
36	176	12.3	618	4	US-09-248-796A-15319	Sequence 19232, A
37	176	12.3	657	4	US-09-248-796A-19232	Sequence 33, Appl
38	176	12.3	848	4	US-09-538-092-93	Sequence 18720, A
39	174.5	12.2	382	4	US-09-248-796A-18720	Sequence 37, Appl
40	173	12.1	1507	4	US-09-914-259-37	Sequence 27827, A
41	171.5	12.0	311	4	US-09-248-796A-27827	Sequence 2, Appl
42	171.5	12.0	903	2	US-08-853-310-2	Sequence 20699, A
43	171	11.9	675	4	US-09-248-796A-20699	Sequence 2, Appl
44	168.5	11.8	667	2	US-08-718-661-2	Sequence 6, Appl
45	168.5	11.8	107	4	US-09-668-119-6	

ALIGNMENTS

RESULT 1
US-08-991-300-2
Sequence 2, Application US/08991300
Patent No. 5973225
GENERAL INFORMATION:
APPLICANT: D'OVIDIO, RENATO
APPLICANT: BORCEDIU, ENRICO
APPLICANT: MERCHITELLI, CINZIA
APPLICANT: CARDELLI, LUISA ERCOLI
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ATTORNEYS
ADDRESS: P.O.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,300
FILING DATE: 16-DEC-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT MI 96/A 002663
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2264-0201-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-300-2

Query Match 32.4%, Score 464, DB 2, Length 369;
Best Local Similarity 40.2%, Pred. No. 9, 2e-35;

RESULT 2
US-08-728-323A-2
Sequence 2, Application US/08728323A
Patent No. 5948676

ADDRESS: Cooper & Dunham Ltd
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.

ZIP: 10035
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A

1 FILING DATE: 435
2 CLASSIFICATION: 435
3 ATTORNEY/AGENT INFORMATION:
4 NAME: White, John P.
5 REGISTRATION NUMBER: 28, 678
6 REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: 212-278-0400
9 TELEFAX: 212-391-0525
10 INFORMATION FOR SEQ ID NO: 2:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 1162 amino acids
13 TYPE: amino acid
14 TOPOLOGY: linear
15 MOLECULE TYPE: protein
16 OS-08-728-523A-2

Query Match	19.18;	Score 273;	DB 2;	Length 1162;
-------------	--------	------------	-------	--------------

RESULT 3
US-09-298-568-2
; Sequence 2, Application US/09298568
; Patent No. 6322792

```

? GENERAL INFORMATION:
? APPLICANT: Kieff, Elliott D.
? APPLICANT: Ballester, Mary E.
? APPLICANT: Kaye, Kenneth M.
? TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
? TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
? FILE REFERENCE: 16412-10001R
? CURRENT APPLICATION NUMBER: US/09/298,568
? CURRENT FILING DATE: 1999-04-21
? EARLIER APPLICATION NUMBER: US 60/109,422
? EARLIER FILING DATE: 1998-11-19
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 2
? LENGTH: 1162
? TYPE: PRN
? ORGANISM: Kaposi's sarcoma-associated herpesvirus
? US-09-298-568-2

```

Query Match	19.1%	Score 273;	DB 3;	Length 1162;
Best Local Similarity	38.4%	Pred. No. 1.2e-16;		
Matches 94; Conservative	19;	Mismatches 100;	Indels 32;	Gaps 10;

[illegible]

RESULT 4

Db 177 VVQQQQQLQQQQQQQHLLKLNHONQQQIQQQQQQLQRIALQLQQQQQQQQQQQQQ-- 234
QY 116 QQLQQQLPCMDVVLQGHNIHARSQVLLQSTYQLLQELCCQHLMQIPESQCCAIHNV 175
Db 235 QQLALEKQ-----PPIQQPMMQPPPPSQALPQ-----QLQOMHT 270
QY 176 VHAIIHQKQKQQQSSQVSPQPLQYR-----LQGSFRPSQ----- 215
Db 271 QH-----HPPPPQPCP--PVAGVPSQLPPSQCTQPLVSAQALPQMLYTPPLKTVRA 324
QY 216 -----QNPQAQSVQPPQLPQFEIRHLALQTLPA 245
Db 325 PMVVQPPVVPQVQVQQQ-----TAVQTAQA 349

RESULT 10
US-09-491-356C-9
Sequence 9, Application US/09491356C
Patent No. 656061
GENERAL INFORMATION:
APPLICANT: Philibert, Robert A.
APPLICANT: Gins, Edward I.
APPLICANT: Delisi, Lynn
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
FILE REFERENCE: 9465, GUS11
CURRENT APPLICATION NUMBER: US/09/491,356C
CURRENT FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: PCT/US99/09365
PRIOR FILING DATE: 1999-04-29
PRIOR APPLICATION NUMBER: 60/083,465
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 2074
TYPE: PRT
ORGANISM: Mus musculus
US-09-491-356C-9

Query Match 15.2%; Score 217; DB 4; Length 2074;
Best Local Similarity 30.6%; Pred. No. 3.2e-11;
Matches 87; Conservative 15; Mismatches 90; Indels 92; Gaps 10;
QY 7 QLPQNPSSQQPQEQVPLVQQQQFPQSQ--QQFPQPPY-----PQPQPPSQQY 55
Db 1807 QQQPVPQGGRLKQ--LQSSQGLQSSVHQWTPSSISGLQTSLSLSLSLSLSLSLSLS 1864
QY 56 LQLQPPFPRLPYRQPSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 89
Db 1865 VGLQHTGADPFRHLQGRSGYVH-QQAPTYGHGLTSTQRFSHQTLQCTPMMGITWPLS 1923
QY 90 -----PQPSISQQAQQQQQQQQQQQQQQQQQQQQQLQLLQQLLPCMDVVLQGHN 135
Db 1924 AGGVQAGVSTSLPESQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 1973
QY 136 IAHARSQVLLQSTYQLLQELCCQHLMQIPESQCCAIHNVHAIIHQKQKQQQSSQV 195
Db 1974 --HRRQQQQQ-----QMLRQQQQQQQQQ-----QQQQQQQQQQQQQQ 2009
QY 196 SFQPPLOQYPLQGSFRPSQNPQAQSVQPPQLPQFEIRHLA 239
Db 2010 QQQQPHQQ-----QQQAPPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 2049

RESULT 11
US-09-491-356C-8
Sequence 8, Application US/09491356C
Patent No. 656061
GENERAL INFORMATION:
APPLICANT: Philibert, Robert A.
APPLICANT: Gins, Edward I.
APPLICANT: Delisi, Lynn
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13

FILE REFERENCE: 9465, GUS11
CURRENT APPLICATION NUMBER: US/09/491,356C
CURRENT FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: PCT/US99/09365
PRIOR FILING DATE: 1999-04-29
PRIOR APPLICATION NUMBER: 60/083,465
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 2023
TYPE: PRT
ORGANISM: Homo sapiens
US-09-491-356C-8

Query Match 14.1%; Score 202; DB 4; Length 2023;
Best Local Similarity 29.2%; Pred. No. 7.4e-10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11;
QY 7 QLPQNPSSQQPQEQVPLVQQQQFPQSQ--QQFPQPPY----- 43
Db 1747 QQQPVPQGGRLKQ--LQSSQGLQSSVHQWTPSSISGLQTSQGYTPVSHVGLQHT 1804
QY 44 ----PQPQPPSQQPYLQLPFPQPLPYR-----QPQSPPPQPPY----- 81
Db 1805 GPAGTWPPSSSQPPYQSTHPTNPVLVDPTRHLQGRSGYVHQAPTYGHGLTSTGRFS 1864
QY 82 ----QPPQVSSQ-----PQPSISQQAQQQQQQQQQQQQQQQQQQQQQLLQ-- 117
Db 1865 HQLQCTPMISTMTMPSAQGVQAGVQAGVQAGVQAGVQAGVQAGVQAGVQAGVQAGV 1924
QY 118 -ILQQQLPCMDVVLQGHNIHARSQVLLQSTYQLLQELCCQHLMQIPESQCCAIHNV 176
Db 1925 HIRQQ-----QQQLRQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 1962
QY 177 HAILHQKQKQ-----QPSQVSPQ--QPLQYPLQGSFRPSQNPQAQSVQPP 227
Db 1963 -----HQQQQQQAPPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 2015

RESULT 12
US-09-538-092-1377
Sequence 1377, Application US/09538092
Patent No. 675314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Tracy A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurPatSeqFormatter Version 0.9
SEQ ID NO 1377
LENGTH: 2124
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number Q93074
US-09-538-092-1377

Query Match 14.1%; Score 202; DB 4; Length 2124;
Best Local Similarity 29.2%; Pred. No. 7.9e-10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11;
QY 7 QLPQNPSSQQPQEQVPLVQQQQFPQSQ--QQFPQPPY----- 43

DB 1847 CQOPAVFGQRLRQ--LQSQGMLGQSSVHQMTPSSVGLQTSQGYTPYVSHVGLQHT 1904
 QY 44 ----PQPPFSSQCPYLQLPFPPRLPVP-----QPSPFPQPPV----- 81
 DB 1905 GPAGTVPFSSSQCPYQSTHPTSTMTLVDPTRHLCQPSGVTHQAPLYHGILSTQFS 1964
 QY 82 ----QPQVQSV-----PQPFISQQAQAAAAQAAAAQAAAAQAAAAQAAAAQAAAA 117
 DB 1965 HQTLLQGTMTSMTPMSAGVAVRSTALPEQAAAAQAAAAQAAAAQAAAAQAAAAQAAAA 2024
 QY 118 -ILQQLIPCMQDVLLQHNIAHARSQVLAQSTYLLGLLCCQHLMQIPESQCCQAHNV 176
 DB 2025 HIRQQQ-----QQQLIRQQAAAAQAAAAQAAAAQAAAAQAAAAQAAAAQAAAA 2062
 QY 177 HAILHQQCKQQQ-----QPSQVSPQ-CPAQYPLGQSGFRPSQSNPQAQGSVQPC 227
 DB 2063 ----HQQQQQQQAAPQPQPQPSQCPQPSQGLQQTQQQQQQAALVRLQQLSLSTQPC 2115

RESULT 13

US-09-248-796A-19253
 Sequence 19253, Application US/09248796A
 Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Helmsstock et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248,796A
 PRIOR FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 19253

LENGTH: 505

TYPE: PRT
 ORGANISM: Candida albicans
 US-09-248-796A-19253

Query Match 13.9%; Score 198.5; DB 4; Length 505;
 Best Local Similarity 27.1%; Pred. No. 2.9e-10;
 Matches 64; Conservative 24; Mismatches 47; Indels 101; Gaps 7;

QY 4 PVP-----QLQPNPSCQCPQCEVPLVQCCQCPFCQCCQCPFPQCPYRQPPFPSSQCPYLQ 59
 DB 365 PVPNNMLBOLQCCQQRQCEKQZ-----KQCEBQQAQQLPQQ----- 401
 QY 60 PFPQRLPFPQSPQSPQCPYRQPPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCP 119
 DB 402 -----EQQSPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCPQCP 451
 QY 120 CQQLIPCMQDVLLQHNIAHARSQVLAQSTYLLGLLCCQHLMQIPESQCCQAHNV 179
 DB 452 QEQ-----PDQCE----- 461
 QY 180 IHHQQCKQQQSSQVSPQCPYLQYPLGSGFRPSQSNPQAQGSVQPCQLPFEFI 235
 DB 462 ----QQCEKQPPQCCQCEKQPPQ-----CPDQHPDQ-----QCEQIQCPENL 503

RESULT 14

US-09-270-767-61220
 Sequence 61220, Application US/09270767
 Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 61220
 LENGTH: 663
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-270-767-61220

Query Match 13.8%; Score 197.5; DB 4; Length 663;
 Best Local Similarity 29.8%; Pred. No. 5e-10;
 Matches 88; Conservative 25; Mismatches 97; Indels 85; Gaps 13;

QY 4 PVPQQLPQNPSCQCPQCEVPLVQCCQCPFCQCCQCPFPQCPYRQPPFPSSQCPYLQ 63
 DB 195 PVAE--EQEQEQEAQRVAVSTQPPQQLQ-----QPPRRQYQSPY-QL-PLP- 242
 QY 64 PRLVPQSPFPQCPYRQPPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCP 105
 DB 243 --LPAPQHSVNPQ-----QQQQQHQPQVPLPEBQFLKLEBELQARAVHQLRQQQHQ 296
 QY 106 CQQQQQQQLLQQLIPCMQDVLLQHNIAHARSQVLAQSTYLLGLLCCQHLMQIPESQCC 154
 DB 297 CQQQQQQQHQQQQQQQH-----HKQLPIHSTH-ATHKVLQADPSLGLGYSRFRVAQE 350
 QY 155 LC-----CQHLWQIPESQCCQAHNVVHAAILHQQCKQQQCPSSQVS 196
 DB 351 LVTPYSHRGGPKYLPPLQAQIQEDDEQQCCQCPQVQLHKPIHPQQLIHGIPQIA 410
 QY 197 PQQPLQYPLGSGFRPSQSNPQAQGSVQPCQLPFEIRNLALQTLPMQCVYI 251
 DB 411 YVQP-----QISYKTLPHNPLAKSLD-----SEIEKLAAKPGQSLAVY 451

RESULT 15

US-09-270-767-45698
 Sequence 45698, Application US/09270767
 Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 PRIOR FILING DATE: 1999-03-17
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 62517
 SEQ ID NO 45698

LENGTH: 1591
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-270-767-45698

Query Match 13.8%; Score 197.5; DB 4; Length 1591;
 Best Local Similarity 29.8%; Pred. No. 1.4e-09;
 Matches 88; Conservative 25; Mismatches 97; Indels 85; Gaps 13;

QY 4 PVPQQLPQNPSCQCPQCEVPLVQCCQCPFCQCCQCPFPQCPYRQPPFPSSQCPYLQ 63
 DB 1123 PVAE--EQEQEQEAQRVAVSTQPPQQLQ-----QPPRRQYQSPY-QL-PLP- 1170
 QY 64 PRLVPQSPFPQCPYRQPPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCPQPSQCP 105
 DB 1171 --LPAPQHSVNPQ-----CQQQQHQQPYVPLPEBQFLKLEBELQARAVHQLRQQQHQ 1224
 QY 106 CQQQQQQQLLQQLIPCMQDVLLQHNIAHARSQVLAQSTYLLGLLCCQHLMQIPESQCC 154
 DB 1225 CQQQQQQQHQQQQQQQH-----HKQLPIHSTH-ATHKVLQADPSLGLGYSRFRVAQE 1278
 QY 155 LC-----CQHLWQIPESQCCQAHNVVHAAILHQQCKQQQCPSSQVS 196
 DB 1279 LVTPYSHRGGPKYLPPLQAQIQEDDEQQCCQCPQVQLHKPIHPQQLIHGIPQIA 1338
 QY 197 PQQPLQYPLGSGFRPSQSNPQAQGSVQPCQLPFEIRNLALQTLPMQCVYI 251
 DB 1339 YVQP-----QISYKTLPHNPLAKSLD-----SEIEKLAAKPGQSLAVY 1379

Wed Dec 15 10:02:01 2004

us-10-089-700-3-r65.rat

Page 7

Search completed: December 14, 2004, 17:29:05
Job time : 20 secs

This Page Blank (uspto)

RESULT 5
US-10-474-955-101

```

; Sequence 101, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:
; APPLICANT: Drifhout, Jan W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludwig, Solid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A
; TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
; FILE REFERENCE: 279/71244-PC-T-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; CURRENT FILING DATE: 2003-10-13
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent version 3.1
; SEQ ID NO 101
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Amino acid sequence of GAMMA-1
US-10-474-955-101

```

```

Query Match          39.0%; Score 558.5; DB 17; Length 282;
Best Local Similarity 47.6%; Pred. No. 4.1e-36;
Matches 138; Conservative 33; Mismatches 72; Indels 47; Gaps 15;

```

```

QY 2 RVHPQL---QPQNSQQQPPQEVPLVQQQPPQ-QQQF---PQQPYPQ-PQPS 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 10 QVMPQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 52 --QQPQYQ--LQPPQPPQ--YPPQSPPPQPPQPPQPPQPPQPPQPPQPPQ 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 69 QPQQTYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 QQQQQQQQILQQQLIPCMQVYLQCHN---IAHARSQVLSQSTYLLQELCCQHLNQ 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 FPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 IPEQSQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL-QQYPLQGSFRPSQQNPQAQ 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 IPQQLCAAIHSIVHSIIMQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 222 GSVQPPQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 230 GIIPQPPQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-10-739-930-9621
; Sequence 9621, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9621
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Trillium aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIA-23APR03-C125_65.P
US-10-739-930-9621

```

```

Query Match          38.3%; Score 547.5; DB 17; Length 298;
Best Local Similarity 46.9%; Pred. No. 3.2e-35;
Matches 134; Conservative 29; Mismatches 74; Indels 49; Gaps 14;
QY 4 PVPQLQPNPQNSQQQPPQEVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQ 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 36 FVP--QPHQPSQPP-----QTFPPQPPQPPHQPQQPPQPPQPPQPPQPP 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 P-QPRLPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 QQQILQQILQQQLIPCMQVYLQCHN---IAHARSQVLSQSTYLLQELCCQHLMOIPEQ 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 141 QPPFIQPSLQQQVNPCKNFKFLQCKPVSLSMSMTMPGSDCQVMQGCCQQLAQIPQ 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 SQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL-QQYPLQGSFRPSQQNPQAQSVQ 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 201 LQCAAIHIVHSIIMQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 PQQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 PQQPQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-10-474-955-99
; Sequence 99, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:
; APPLICANT: Drifhout, Jan W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludwig, Solid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
; TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
; FILE REFERENCE: 279/71244-PC-T-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; CURRENT FILING DATE: 2003-10-13
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent version 3.1
; SEQ ID NO 99
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Amino acid sequence of GAMMA-4
US-10-474-955-99

```

```

Query Match          37.0%; Score 529; DB 17; Length 279;
Best Local Similarity 46.3%; Pred. No. 8.6e-34;
Matches 133; Conservative 26; Mismatches 75; Indels 52; Gaps 14;

```

```

QY 4 PVPQLQPNPQNSQQQPPQEVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQ 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 18 FVP--QPHQPSQPP-----QTFPPQPPQPPHQPQQPPQPPQPPQPPQPP 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 P-QPRLPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 QQQILQQILQQQLIPCMQVYLQCHN---IAHARSQVLSQSTYLLQELCCQHLMOIPEQ 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 123 QPPFIQPSLQQQVNPCKNFKFLQCKPVSLSMSMTMPGSDCQVMQGCCQQLAQIPQ 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 SQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL-QQYPLQGSFRPSQQNPQAQSVQ 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 183 LQCAAIHIVHSIIMQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 PQQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 PQQPQLPQFEIRNLALQTLPMQNVYIAPYCTI--APF-----GIFG 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 8
US-10-474-955-97
; Sequence 97, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:

```

APPLICANT: Drifhout, Jan W.
 APPLICANT: Konig, Frits
 APPLICANT: McAdam, Stephan N.
 APPLICANT: Ludwig, Solid Magne
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A
 FILE REFERENCE: 2799/71244-PCT-US
 CURRENT APPLICATION NUMBER: US/10/474,955
 CURRENT FILING DATE: 2003-10-13
 NUMBER OF SEQ ID NOS: 137
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 97
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Consensus amino acid sequence
 SS-10-474-955-97

Query Match 36.7%; Score 525; DB 17; Length 279;

Best Local Similarity 46.2%; Pred. No. 1.8e-33; Indels 52; Gaps 14;

Matches 132; Conservative 27; Mismatches 75;

DB 4 PVPOLQPNPSQQQPEQVPLVQQQQPFQGGQQQPF--PQQPYPOQPFPSQGYLQLPF 61
 DB 18 PVP--QPHQPFSSQPF-----QTFPQPPQTFPHQPPQGGPQPF--PQQPYLQPPQPF 66
 DB 62 P-QPRLPYPOQPFPSQPYPPQPF-----PQYSPQPPHISQGGQPF-----QGGGQQGGGQ 110
 DB 67 PQQPQQPYPPQPF-----PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122
 DB 111 QQQLLQQQLLQQLLPCMDVVLQGHN-----IAHARSQVLSQSTYQLLQELCCQHLMOIPEQ 166
 DB 123 QPFFIQPSLQQQVNPCKNPLLQCKKFSVLSVSMIMWPSQSCQVWRQSSQGLAQIPQ 182
 DB 167 SQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 225
 DB 183 LQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 225
 DB 226 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276
 DB 231 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276

RESULT 9

US-10-474-955-98

Sequence 98, Application US/10474955

Publication No. US20040241161A1

GENERAL INFORMATION:

APPLICANT: Drifhout, Jan W.

APPLICANT: Konig, Frits

APPLICANT: McAdam, Stephan N.

APPLICANT: Ludwig, Solid Magne

TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A

TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A

FILE REFERENCE: 2799/71244-PCT-US

CURRENT APPLICATION NUMBER: US/10/474,955

CURRENT FILING DATE: 2003-10-13

NUMBER OF SEQ ID NOS: 137

SOFTWARE: PatentIn version 3.1

SEQ ID NO 98

LENGTH: 279

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Amino acid sequence of GAMMA-2

US-10-474-955-98

Query Match

36.4%; Score 521; DB 17; Length 279;

Best Local Similarity 45.8%; Pred. No. 3.7e-33;

Matches 131; Conservative 27; Mismatches 76; Indels 52; Gaps 14;

4 PVPOLQPNPSQQQPEQVPLVQQQQPFQGGQQQPF--PQQPYPOQPFPSQGYLQLPF 61

DB 18 PVP--QPHQPFSSQPF-----QTFPQPPQTFPHQPPQGGPQPF--PQQPYLQPPQPF 66
 DB 62 P-QPRLPYPOQPFPSQPYPPQPF-----PQYSPQPPHISQGGQPF-----QGGGQQGGGQ 110
 DB 67 PQQPQQPYPPQPF-----PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122
 DB 111 QQQLLQQQLLQQLLPCMDVVLQGHN-----IAHARSQVLSQSTYQLLQELCCQHLMOIPEQ 166
 DB 123 QPFFIQPSLQQQVNPCKNPLLQCKKFSVLSVSMIMWPSQSCQVWRQSSQGLAQIPQ 182
 DB 167 SQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 225
 DB 183 LQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 225
 DB 226 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276
 DB 231 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276

RESULT 10

US-10-474-955-100

Sequence 100, Application US/10474955

Publication No. US20040241161A1

GENERAL INFORMATION:

APPLICANT: Drifhout, Jan W.

APPLICANT: Konig, Frits

APPLICANT: McAdam, Stephan N.

APPLICANT: Ludwig, Solid Magne

TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTOR

TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTOR

FILE REFERENCE: 2799/71244-PCT-US

CURRENT APPLICATION NUMBER: US/10/474,955

CURRENT FILING DATE: 2003-10-13

NUMBER OF SEQ ID NOS: 137

SOFTWARE: PatentIn version 3.1

SEQ ID NO 100

LENGTH: 279

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Amino acid sequence of GAMMA-3

US-10-474-955-100

Query Match

36.4%; Score 521; DB 17; Length 279;

Best Local Similarity 46.0%; Pred. No. 3.7e-33;

Matches 131; Conservative 29; Mismatches 75; Indels 50; Gaps 14;

DB 4 PVPOLQPNPSQQQPEQVPLVQQQQPFQGGQQQPF--PQQPYPOQPFPSQGYLQLPF 61
 DB 18 PVP--QPHQPFSSQPF-----QTFPQPPQTFPHQPPQGGPQPF--PQQPYLQPPQPF 66
 DB 62 P-QPRLPYPOQPF-----PQYSPQPPHISQGGQPF-----QGGGQQGGGQ 110
 DB 67 PQQPQQPYPPQPF-----PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 123
 DB 111 QQQLLQQQLLQQLLPCMDVVLQGHN-----IAHARSQVLSQSTYQLLQELCCQHLMOIPEQ 167
 DB 123 QPFFIQPSLQQQVNPCKNPLLQCKKFSVLSVSMIMWPSQSCQVWRQSSQGLAQIPQ 183
 DB 167 SQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 226
 DB 184 LQCAIHNVVAIILHQQQKQQQSSQVSTFQPL--QYPLGGSSFRPSQNPQAGSVQ 226
 DB 227 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276
 DB 232 PQQPQLEAIRSLVLTPLTMCNVVVPFECGSIIRAPFSSVAGIGG 276

RESULT 11

US-10-739-930-9778

Sequence 9778, Application US/10739930

Publication No. US20040216190A1

```

Query Match Similarity 30.8%; Score 440.5; DB 17; Length 307;
Best Local Similarity 42.8%; Pred. No. 9,1.e-27;
Matches 125; Conservative 37; Mismatches 75; Indels 55; Gaps 16;

QY 13 PSQQQPOQVPLVQQQQP-----GQQQQFPPOQDPYPOQPFPSQQPYLQQLPFPQPLP 67
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 27 PGLRPMWQQQPLPPOQTFPQQLFLSQQQQ---QLQLFPQPSFSQQP-----PFWQQQPP 78
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 68 YPQPSFPQQP-----YPQSPQYSQPPQPI---SQQAQAQQQQQQQQQQQQQQQI 114
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 79 FSQQQPLPQQPFPSQQQQLVLPQ-QPFSQQQQLVLPQPSFPQSQQQQLQLVQQQIP 137
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 14
US-10-425-115-200100
Sequence 200100, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(5322)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:02:00 ; Search time 15 Seconds
(without alignments)
1706.244 Million cell updates/sec

Title: US-10-089-700-3-R65

Perfect score: 1431

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1364	95.3	286	2	S07923	alpha/beta-gliadin
2	1356	94.8	286	1	EMWTA	alpha/beta-gliadin
3	1312	91.7	288	2	T06282	alpha-gliadin prec
4	1267.5	88.6	291	2	T06488	alpha/beta-gliadin
5	1267.5	88.6	307	2	S10015	alpha/beta-gliadin
6	1263	88.3	296	2	S07363	alpha/beta-gliadin
7	1261	88.1	313	2	S07924	alpha/beta-gliadin
8	1257	87.8	296	2	A27319	gliadin - wheat
9	1248.5	87.2	319	2	A22364	alpha/beta-gliadin
10	1245.5	87.0	319	2	C22364	alpha/beta-gliadin
11	1235.5	86.3	297	2	T06500	alpha/beta-gliadin
12	1223	85.5	326	2	D22364	alpha/beta-gliadin
13	1206	84.3	320	2	E22364	alpha/beta-gliadin
14	1183	82.7	282	2	T06504	alpha/beta-gliadin
15	1155.5	80.7	292	2	B22364	alpha/beta-gliadin
16	589	41.2	327	2	T50402	gamma-gliadin prec
17	549.5	38.4	291	1	EMWTA	gamma-gliadin B p
18	541	37.8	302	2	TA0153	gamma-gliadin prec
19	486	34.0	305	2	S08312	gamma-hordein 1 p
20	469	32.8	286	2	T05718	gamma-hordein 3 -
21	461.5	32.0	374	2	S05923	glutenin low molec
22	457.5	32.0	359	2	T06982	glutenin low molec
23	454.5	31.8	356	2	S01992	glutenin low molec
24	447.5	31.3	304	2	T06505	glutenin low molec
25	446	31.2	293	2	S07365	hordein B1 - barley
26	445	31.1	264	2	S07975	B3-hordein (clone
27	443	31.0	286	2	T05910	glutenin low molec
28	442	30.9	285	2	S05863	glutenin low molec
29	441.5	30.9	298	2	T06980	glutenin low molec

30	504322	2	S04322	glutennin low molec
31	440.5	30.8	307 2	low-molecular-weight
32	439.5	30.7	290 2	hordein B precursor
33	428	29.9	251 2	gamma-gliadin prec
34	426.5	29.8	271 2	T04474
35	417	29.1	295 1	glutenin low molec
36	410.5	28.7	295 2	avenin precursor
37	407	28.4	220 2	avenin precursor (
38	391.5	27.4	194 2	gamma-secalin - ry
39	385	26.9	243 2	hordein (clone
40	384	26.8	244 2	gamma-gliadin B pr
41	382.5	26.7	222 2	avenin precursor (
42	377	26.3	261 2	glutenin low molec
43	376	26.3	182 2	avenin N9 - oat
44	374.5	26.2	323 2	glutenin low molec
45	367.5	25.7	276 2	glutenin low molec

ALIGNMENTS

RESULT 1

alpha/beta-gliadin precursor - wheat
A:Species: Triticum aestivum (common wheat)
C:Date: 08-Jun-1994 #sequence_revision 01-Dec-1995 #text_change 20-Aug-1999
C:Accession: S07923
R:Summer-Smith, M.; Rafaleki, J.A.; Sugiyama, T.; Scoll, M.; Soell, D.
Nucleic Acids Res. 13, 3905-3916, 1985
A:Title: Conservation and variability of wheat alpha/beta-gliadin genes.
A:Reference number: S07361; PMID:85242077; PMID:3833304
A:Accession: S07923
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1286 <SUM>
A:Cross-references: EMBL:X02539, NID:g21760, PIDN:CAA2384.1, PID:g21761
A:Superfamily: gliadin

Query Match	95.3%;	Score 1364;	DB 2;	Length 286;
Best Local Similarity	95.9%;	Pred. No. 6.9e-84;		
Matches 255;	Conservative 1;	Mismatches 10;	Indels 0;	Gaps 0;

Qy	1	VRRPVQLQGNQNSQQQ	PQEQVPLVWQQQ	PFQGGQQQ	PPRQCPYRQ	PPRQSPRYLQ	Q	60	
Dp	21	VRRFVPLQGNQNSQQQ	PQEQVPLVWQQQ	PLGQQCP	PPRQCPYRQ	PPRQSPRYLQ	Q	80	
Qy	61	FPQPRLEYPQ	QSPFPQ	QPYRQ	PQYRQ	PSQ	QPI	120	
Dp	81	FPQGLFYSQ	QGFRRQ	QCPYRQ	PPQYRQ	PSQ	QPI	140	
Qy	121	QQLIPCMDVYLQ	QCHNTHAHSQ	VLQSTYQLQ	LOELCCCHLMQ	IPESQCCAL	HNHVHAT	180	
Dp	141	QQLIPCMDVYLQ	QCHNTHAHSQ	VLQSTYQLQ	LOELCCCHLMQ	IPESQCCAL	HNHVHAT	200	
Qy	181	LHQOQKQOQ	QSSQVS	FQOPLQ	QYPLGQGS	FRPSQCN	PPAQGS	VPDQL	240
Dp	201	LHQOQKQOQ	QSSQVS	FQOPLQ	QYPLGQGS	FRPSQCN	PPAQGS	VPDQL	260
Qy	241	QTLPAKCNVYI	APYCTI	APRGIGTN					266
Dp	261	QTLPAKCNVYI	APYCTI	APRGIGTN					286

RESULT 2

alpha/beta-gliadin precursor - wheat
N/Alternate names: prolamins
C/Species: *Triticum aestivum* (common wheat)
C/Date: 28-May-1966 #sequence_revision 28-May-1966 #text_change 09-Jul-2004
C/Accession: A03354
E/Rabalaki, J.A.; Schneers, K.; Metzlery, M.; Peterson, D.M.; Hedgcock, C.; Soll, D.G.
EMBO J. 3, 1409-1415, 1984
#Title: Developmentally regulated plant genes: the nucleotide sequence of a wheat gliadin

QY 112 QGIIQQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 171
 DB 141 QGIIQQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 200
 QY 172 IHHVHAIIHQQQKQSSQVSPQPLQGYPLQGGSPSPQNPQAGSVQPLQ 231
 DB 201 IHHVHAIIHQQQKQSSQVSPQPLQGYPLQGGSPSPQNPQAGSVQPLQ 259
 QY 232 FEETRNALQTLPMQVYIAPYC--TIAPFGIGTN 266
 DB 260 FEETRNALQTLPMQVYIAPYC--TIAPFGIGTN 296

RESULT 9

A22364
 alpha/beta-gliadin precursor (clone A42) - wheat
 A:Alternate names: prolamum
 C:Species: Triticum aestivum (common wheat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
 C:Accession: A22364
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
 J: Biol. Chem. 260, 8203-8213, 1985
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
 A:Reference number: A92541; MUID:85234522; PMID:2989281
 A:Accession: A22364
 A:Molecule type: mRNA
 A:Residues: 1-319 <OKI>
 A:Cross-references: UNIPROT:P04725; GB:M11073; NID:g170715; PIDN:AAA4278.1; PID:g170716
 C:Superfamily: gliadin
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-319/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 87.2%; Score 1248.5; DB 2; Length 319;
 Best Local Similarity 81.9%; Pred. No. 3.6e-76;
 Matches 245; Conservative 6; Mismatches 15; Indels 33; Gaps 4;

QY 1 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 60
 DB 21 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 80
 QY 61 FPGQ-----PRLPYPPQSPFPQGPYPQPPSPQGYLQLP 114
 DB 81 FPGQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFP 140
 QY 115 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 174
 DB 141 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 200
 QY 175 VHAIIH-----QQKQGGQSSQVSPQPLQGYPLQGG 209
 DB 201 VHAIIHQQQKQSSQVSPQPLQGYPLQGGSPSPQNPQAGSVQPLQ 260
 QY 210 SFRPQNPQAGSVQPLQGYPLQGGSPSPQNPQAGSVQPLQ 266
 DB 261 SFRPQNPQAGSVQPLQGYPLQGGSPSPQNPQAGSVQPLQ 319

RESULT 10

C22364
 alpha/beta-gliadin precursor (clone A212) - wheat
 C:Species: Triticum aestivum (common wheat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
 C:Accession: C22364
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
 J: Biol. Chem. 260, 8203-8213, 1985
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
 A:Reference number: A92541; MUID:85234522; PMID:2989281
 A:Accession: C22364
 A:Molecule type: mRNA
 A:Residues: 1-319 <OKI>
 A:Cross-references: UNIPROT:P04722
 C:Superfamily: gliadin
 F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-319/Product: alpha/beta-gliadin #status predicted <MAT>
 Query Match 87.0%; Score 1245.5; DB 2; Length 319;
 Best Local Similarity 80.9%; Pred. No. 5.7e-76;
 Matches 242; Conservative 8; Mismatches 16; Indels 33; Gaps 4;

QY 1 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 60
 DB 21 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 80
 QY 61 FPGQ-----PRLPYPPQSPFPQGPYPQPPSPQGYLQLP 114
 DB 81 FPGQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFP 140
 QY 115 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 174
 DB 141 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 200
 QY 175 VHAIIH-----QQKQGGQSSQVSPQPLQGYPLQGG 209
 DB 201 VHAIIHQQQKQSSQVSPQPLQGYPLQGGSPSPQNPQAGSVQPLQ 260
 QY 210 SFRPQNPQAGSVQPLQGYPLQGGSPSPQNPQAGSVQPLQ 266
 DB 261 SFRPQNPQAGSVQPLQGYPLQGGSPSPQNPQAGSVQPLQ 319

RESULT 11

T06500
 alpha/beta-gliadin A-IV precursor - wheat
 C:Species: Triticum aestivum (common wheat)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T06500
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
 J: Biol. Chem. 260, 8203-8213, 1985
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
 A:Reference number: A92541; MUID:85234522; PMID:2989281
 A:Accession: T06500
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-257 <OKI>
 A:Cross-references: UNIPROT:P04724; EMBL:M11075; NID:g170723; PIDN:AAA4282.1; PID:g170724
 C:Superfamily: gliadin
 C:Keywords: seed; storage protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-257/Product: alpha/beta-gliadin A-IV #status predicted <MAT>

Query Match 86.3%; Score 1235.5; DB 2; Length 297;
 Best Local Similarity 84.3%; Pred. No. 2.4e-75;
 Matches 236; Conservative 10; Mismatches 17; Indels 17; Gaps 3;

QY 1 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 60
 DB 21 VRVPVQLQPNPSSQQLQGVPLVQGGQFPGQGGQFPGQGPYPQPPSPQGYLQLP 80
 QY 61 F-----PRLPYPPQSPFPQGPYPQPPSPQGYLQLP 113
 DB 81 FPGQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFPQPPFP 140
 QY 114 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 173
 DB 138 LQQLIQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 200
 QY 174 VHAIIH-----QQKQGGQSSQVSPQPLQGYPLQGG 209
 DB 198 NVVHAIIHQQQKQSSQVSPQPLQGYPLQGGSPSPQNPQAGSVQPLQ 257
 QY 227 QGIIQQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 266
 DB 258 QGIIQQQLIPCMQVVLQGHNIHARSQVLAQSTYQLQLQELCCQHLWQIPESQCA 297

RESULT 12

A:Residues: 1-292 <OKI>
A:Cross-references: UNIPROT:P04721
C:Superfamily: gliadin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-292/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 80.7%; Score 115.5; DB 2; Length 292;
Best local Similarity 84.1%; Pred. No. 4.9e-70;
Matches 227; Conservative 3; Mismatches 9; Indels 31; Gaps 4;

```
QY 1 VRVPYPOLQPNPSQQQPFQEQVPLVQQQQQFPGQQQQQFPFQQPYPFQQPQFPFSSQCPYLQLQP 60
DB 21 VRVPYPOLQPNPSQQQPFQEQVPLVQQQQQFPGQQQQQFPFQQPYPFQQPQFPFSSQCPYLQLQP 80
QY 61 F--PQ---PRLPYQPGSFPFQQPYPFQQPQYSPQPPIS-QQQAQQQQQQQQQQQQQQQQQ 114
DB 81 FLQPPFPFPPQLPYSQPFQPFPPFQQPYPFQQPQYSPQPPISQQQQQQQQQQQQQQQQQQ 140
QY 115 LQQLLQQQLLPCMDVVLQGHNTAHARSGVLQGSTYQLLQELCCQHLWQIPQSSQCAIHN 174
DB 141 IQQLLQQQLLPCMDVVLQGHNTAHARSGVLQGSTYQLLQELCCQHLWQIPQSSQCAIHN 200
QY 175 VVHAIIILH-----QQQKQQQQPSSQSVSPQCFLLQYPLGGG 209
DB 201 VVHAIIILHQQQQQQQEQKQLQQQQQQQQQLQQQQQQKQQQQPSSQSVSPQCFLLQYPLGGG 260
QY 210 SFRPSQNPQAQGSYQPPQLPQFERINTLA 239
DB 261 SFRPSQNPQAQGSVQPPQLPQFERINTLA 290
```

Search completed: December 14, 2004, 17:27:03
Job time : 16 secs

Db 242 QTLPAMCNVYIPRYCTIAFGIFGTN 267

Query	DB	Score	Length	Mismatches	Indels	Gaps
1	VRPVPOLQPNPNSQQQPOEOYELVQOQOFPBQQQQFPPOQYPPPOPPQPFPSQPYLQLOP	94.4%	274	1	5	1
2	VRTVVPOLQPNDSQQQPOEOYELVQOQOFTLQOQOFPPOQYPPPOPPQPFPSQPYLQLOP	94.1%	274	1	5	1
3	FPQRLPFPQPSFPPOQYPPPOPPQPSQPPIS----	94.1%	274	1	5	1
4	FPQRLPFPQPSFPPOQYPPPOPPQPSQPPIS----	94.1%	274	1	5	1
5	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
6	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
7	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
8	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
9	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
10	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
11	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
12	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
13	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
14	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
15	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
16	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
17	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
18	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
19	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
20	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
21	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
22	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
23	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
24	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
25	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
26	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
27	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
28	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
29	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
30	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
31	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
32	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
33	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
34	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
35	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
36	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
37	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
38	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
39	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
40	RNALQTLPLAMCNVYIAPYCTIAPGIFGTN	94.1%	274	1	5	1
41	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
42	QQLLQQLLPCMDVVLQCNINHAASQVLCSTYLLDELCCQHMQIPBSQCCALHNV	94.1%	274	1	5	1
43	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1
44	VHAIIIIHQOQKQOQSSQVSPFOQPLQOQFPLQOQSFPSQONPQAOQSVQOQLPQFEI	94.1%	274	1	5	1

SQ	SEQUENCE	276 AA;	32213 MW;	6A2CCDC4E70BA100A CRC64;
Query Match	94.4%;	Score 1350.5;	DB 2;	Length 276;
Best Local Similarity	93.4%;	Pred. No. 1.4e-74;		
Matches 255; Conservative	2;	Mismatches 9;	Indels 7;	Gaps 1;
OY	1 VRVVPQLQPQNPSQQCPQECPVLVQQGQFPFGQQQPFPPQGPYPQGPFPFSQQPYLQLQP	60		
DH	2 VRVVPQLQPQNPSQQCPQECPVLVQQGQFLGGQGFPPFQGPYPQGPFPFSQQPYLQLQP	61		
OY	61 PPQPLRPYPQGSFFPQGPYPQGPQYSQGPGPISSQQQAQQQQQQQQQQ-----QQ	113		
DH	62 PPOPLRSLSPQPPFPQGPYPQGPQYSQPQPISQQQQQQQQQQQQQQQQQQQQQQDEQ	121		
OY	114 ILQGLILQQGLIFCMWVLIQHNIHAARSQVLAQSIVYLLELCCHLMQIRESGCQAIIH	173		
DH	122 ILQGLILQQGLIIPCMWVLIQHNIHAARSQVLAQSIVYLLELCCHLMQIRESGCQAIIH	181		
OY	174 NVNVAITTHHQCKQQQSQSQQVSFPQPLIQVYLGGSRRPSQQNQDAQSVQPCQLPGE	233		
DH	182 NVNVAITTHHQCKQQQSQSQQVSFPQPLIQVYLGGSRRPSQQNPDAQSVQPCQLPGE	241		
OY	234 ETRNLALOTLPAMCNVIAPYCTIAPEGIPTIN	266		
DH	242 ETRNLALOTLPAMCNVIAPYCTMAPEGIPTIN	274		
RESULT 7				
Q41509	PRELIMINARY;	PRT:	287 AA.	
ID Q41509				
AC Q41509;				
DT 01-NOV-1996	(TREMELREL. 01, Created)			
DI 01-NOV-1996	(TREMELREL. 01, Last sequence update)			
DL 01-MAR-2004	(TREMELREL. 26, Last annotation update)			
DR Alpha-Gladin.				
OS Triticum aestivum (Wheat).				
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;				
OC Triticeae; Triticum.				
OX NCBI_TaxId=4565;				
RN (1)				
RF SEQUENCE FROM N.A.				
RC STRAIN=Cheyenne;				
RA Blechl A.E., Anderson O.D.;				
RL Submitted (Apr-1994) to the EMBL/GenBank/DDBJ databases.				
RS EMBL: U08287; AAA17741.1; -.				
DR PIR: S1333; S1333.				
DR GO: G0045735; F.nutrient reservoir activity; IEA.				
DR InterPro: IPR003612; AA1.				
DR InterPro: IPR001376; Gliadin.				
DR InterPro: IPR001954; Gliadin-glutenin.				
DR Pfam: PF00234; TRYX_alpha_amy1; 1.				
DR PRINTS: PR00208; GLIADGLUTEN.				
DR PRINTS: PR00209; GLIADIN.				
DR SMART: SM00499; AA1; 1.				
SQ SEQUENCE 287 AA; 33047 MW; EF960C21CFB84873 CRC64;				
Query Match	94.0%;	Score 1345.5;	DB 2;	Length 287;
Best Local Similarity	94.8%;	Pred. No. 3e-74;		
Matches 253; Conservative	3;	Mismatches 10;	Indels 1;	Gaps 1
OY	1 VVVVPVLQLOPNMSQQCPQEQAVLVQQGQFPFGQQQFPFPQGPYPQGPFPFSQQPYLQLQP	60		
DH	21 VVRVPHQLQOPNPSQQQPOSQVFLVQQGQFLGGQGFPPFQGPYPQGPFPFSQQPYLQLQP	60		
OY	61 FPQPLRPYPQGSFFPQGPYPQGPQYSQPQPIS-QQQAQQQQQQQQQQQQQQQQLLIQLIL	119		
DH	81 FLQPLQPLSPGPFRRPQGPYPQGPQYSQPQPISQQQQQQQQQQQQQQQQQQQQQILLIQLIL	140		
OY	120 QQGLICMNVVLIQHNIHAARSQVLAQSIVYLLELCCHLMQIRESGCQAIIHVNAI	179		
DH	141 QQGLICMNVVLIQHNIHEKSYLAQSIVYLLELCCHLMQIRESGCQAIIHVNAI	200		

QY	180	ILH00XQ0000SSQVSVFQ0PLQ0YPLQGSGFRPSQGNPDAQSGVQ0LQ0F0FEIINLA	233
Db	201	ILH00XQ0000PSSQVSVFQ0PLQ0YPLQGSGFRPSQGNPDAQSGVQ0LQ0F0FEIINLA	260
QY	240	LQTLPAMCNVYIADYCTIAPPGIFGTN	266
Db	261	LQTLPAMCNVYIADYCTIAPPGIFGTN	287
RESULT 8			
ID	Q9M4L8	PRELIMINARY;	PRT; 277 AA.
AC	Q9M4L8;		
DT	01-OCT-2000 (TREMBlrel. 15, Created)		
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)		
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)		
DE	Alpha-glialdin.		
OS	Triticum aestivum (Wheat).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae;		
OC	Triticeae; Triticum.		
CX	NCBI_Taxid=4565;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MoJeller; TISSUE=Endosperm;		
RA	Arentz-Hansen E.H., Wcadam S.N., MoJberg O., Kristiansen C.,		
RA	Sollid L.M.;		
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.		
DR	CGI: A1133610; CAB76962.1; -.		
DR	GO: GO:0045735; F:nutrient reservoir activity; IEA.		
DR	InterPro; IPR003612; AAI.		
DR	InterPro; IPR001376; Gladin.		
DR	InterPro; IPR001954; Glia_gluenin.		
DR	Pfam; PF00234; TRYP_alphaAmy; 1.		
DR	PRINTS; PR00208; GLIADGLUTEN.		
DR	PRINTS; PR00209; GLIADIN.		
DR	SMART; SMO0499; AAI; 1.		
FT	CHAIN 1 277 alpha-glialdin.		
SQ	SEQUENCE 277 AA; 32371 MW; 73DB85D815E5329D CRC64;		
Query Match			
	Best Local Similarity	93.9%; Score 1344; DB 2; Length 277;	
	Matches 254; Conservative 1; Mismatches 11; Indels 8; Gaps 1		
QY	1	VRVVPICQIQPNPSSQCFQ0FQ0VPLVQ0Q0FPQ0Q0Q0FPQ0PQYPPQ0PFPSSQ0PYLQ0P	60
Db	2	VRVPPIQIQPNPSSQCFQ0FQ0VPLVQ0Q0FTLQ0Q0SFPQ0PFPQ0PFPSSQ0PYLQ0P	61
QY	61	FPQ0PLPQ0PQ0SFPQ0PQ0PYPPQ0PQ0YQ0PQ0PIS-----QQAQ0Q0Q0Q0Q0Q0Q0	112
Db	62	FPQ0PLPQ0PQ0PQ0PQ0PYPPQ0PQ0YQ0PQ0PISQ0Q0Q0Q0Q0Q0Q0Q0Q0Q0Q0Q0Q0	121
QY	113	QILQ0ILQ0QILPCMDVVLQGHNTHASQVYQSTYQLLQELCCGHMQRPESSQCCAI	172
Db	122	QILQ0ILQ0QILPCMDVVLQGHNTHAGRSSQVYQSTYQLLQELCCGHMQLPESQCCAI	181
QY	173	HNHVAHIIH00Q0K00Q0PSSQVSVFQ0PLQ0YPLQGSGFRPSQGNPDAQSGVQ0LQ0F	232
Db	182	HNHVAHIIH00Q0K00Q0PSSQVSVFQ0PLQ0YPLQGSGFRPSQGNPDAQSGVQ0LQ0F	241
QY	233	EEIRNIALQTLPAMCNVYIADYCTIAPPGIFGTN	266
Db	242	EEIRNIALQTLPAMCNVYIADYCTIAPPGIFGTN	275
RESULT 9			
ID	Q9M4M0	PRELIMINARY;	PRT; 276 AA.
AC	Q9M4M0;		
DT	01-OCT-2000 (TREMBlrel. 15, Created)		
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)		
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)		
DE	Alpha-glialdin.		

OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Focaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mjcelner; TISSUE=Endosperm;
RA Arezt-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,
RA Sollid B.W.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ133608; CAB76960.1; -
DR GO; GO:0045735; P:nutrient reservoir activity; IEA.
DR InterPro; IPR003642; AAI.
DR InterPro; IPR001376; Gliadin.
DR Pfam; PF00234; TRYF_alpha_glu1n.
DR PRINTS; PR00208; GLIADIN.
DR SMART; SM00499; AAI; 1.
DR CHAIN 1 276 alpha-gliadin.
SQ SEQUENCE 276 AA; 32211 MW; 6A3E9723D42B100A CRC64;

Query Match 93.7%; Score 1341.5; DB 2; Length 276;
Best Local Similarity 92.7%; Pred. No. 5, 1e-74;
Matches 253; Conservative 3; Mismatches 10; Indels 7; Gaps 1

CY 1 VRVVPOLQPCNRSSQQPQEGLVPLVQOQQPFPGQGGQFPPOGYPPOPSPSGPYLQLQP 60
DB 2 VRVVPQLQPCNPSSQQPQEGLVPLVQOQQFLGQQQFPFQGYPOPDPSSQPYLQLQP 61
CY 61 FPQPLPFPQPSFPPQGPPOPQPSQPPQSQQQQAQQQQQQQQ-----CQ 113
DB 62 FPQPLPSQPPGFPPQGPPOPQPSQPPQSQQQQAQQQQQQQQQQQQEQ 121
CY 114 ILQQLLQQLLCMDVVLQGNHIAHRSVYLQOSTYLLQLLECCGHMIPQSSCGAH 173
DB 122 ILQQLLQQLLPQMNVVLQGNHIAHRSVYLQOSTYLLQLLECCGHMIPQSSCGAH 181
CY 174 NVVAHIIHQCKKQCKQSSQVSFGCPLOQYPLGGGSFSPSQCNPDAGSVPPQLPPE 233
DB 182 NVVAHIIHQCKKQCKQSSQVSFGCPLOQYPLGGGSFSPSQCNPDAGSVPPQLPPE 241
CY 234 EIRNIALQTLPMKNVVIAPYCTTAPFGIFGTN 266
DB 242 EIRNIALQTLPMKNVVIAPYCTTAPFGIFGTN 274

RESULT 10
Q41531 PRELIMINARY; PRT; 289 AA.

ID Q41531
AC Q41531
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DS Alpha-glialdin storage protein.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
CX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cheyenne;
RA Anderson O.D.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U51306; AAA96524.1; -
DR PIR; S1333; S1333.
DR GO; GO:0045735; P:nutrient reservoir activity; IEA.
DR InterPro; IPR003642; AAI.
DR InterPro; IPR001376; Gliadin.
DR InterPro; IPR001954; Glia_glutenin.
DR Pfam; PF00234; TRYF_alpha_glu1; 1

DR	PRINTS; PR00208; GLIADJLUTEN.
DR	PRINTS; PR00209; GLIADIN.
DR	SMART; SMO0489; AAI ¹ .
SO	SEQUENCE 289 AA; 33549 MW; 5F577C9CD63874FA CXC64;
 Query Match	
Best Local Similarity 93.5%; Score 1337.5; DB 2; Length 289;	
Matches 253; Conservative 2; Mismatches 11; Indels 3; Gaps 1	
OY	1 VRVVPOLQPNPNSQQQPQSEVPLVVOOQFFGGOQQQPPFPQPYRQPFQFBSQQPRLQLQ 60
DB	21 VRVVPOLQPNPNSQQQPQSEVPLVVOOQFFGGOQQPPFPQPYRQPFQFBSQQPRLQLQ 80
OY	61 PPGRLPFPQSPSPPOQPYRQPFQPSQPGQPIIS---QQAAQQQQQQQQQQQQQQQQQLIQ 117
DB	81 PPGRLPFPQSPSPPRPQPYRQPFQPSQPGQPIISQQQQQQQQQQQQQQQQQQQQQLIQ 140
OY	118 ILQGQLIPCMDVLIQHNIHAHSQVTQQSTYYLLQBLCCGHLMQIPBQSCCAHHNVH 177
DB	141 ILQGQLIPCMDVVLQHNIHVGRSQVYQGSTYQLLRHLCCGHLMQIPBQSCCAHHNVH 200
OY	178 AIILHQOQKQOQSSQSVSPQGPLOXYPLGGSGFRSPSQGNPAQGSVPQQLPPEETR 237
DB	201 AIILHQOQKQOQSSQSVSPFGPQLQYPLGGSGFRSPSQGNPAQGSVPQQLPPEETR 260
OY	238 LALQTLPAMCNVIAPFYCTIAFPGFIGN 266
DB	261 LALQTLPAMCNVIIPYCTIAFPGFIGN 289
 RESULT 11	
ID	O9M4L9 PRELIMINARY; PROT; 270 AA.
AC	O9M4L9;
DT	01-OCT-2000 (TReMBRel. 15. Created)
DT	01-OCT-2000 (TReMBRel. 15. Last sequence update)
DT	01-MAR-2004 (TReMBRel. 26. Last annotation update)
DE	Alpha-gliadin.
OS	Triticum aestivum (Wheat).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
CX	NCBI_TaxID=4565;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Mjoelner; TISSUE=Endosperm;
RA	Arentz-Hansen E.H., Mcadam S.N., Molberg O., Christiansen C.,
RA	Solid LM;
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
EMBL	AU133609; CAB76961.1; -
GO	GO:0045735; F:nutrient reservoir activity; IEA.
DR	InferPro; IPR003612; AAI.
DR	InferPro; IPR001376; Gliadin.
DR	InferPro; IPR001954; Glia glutenin.
Pfam	PF00234; Tryp_alpha_amy1; 1.
DR	PRINTS; PR00208; GLIADJLUTEN.
DR	PRINTS; PR00209; GLIADIN.
DR	SMART; SMO0499; AAI ¹ .
FT	CHAIN 1 270 alpha-gliadin.
SEQ	SEQUENCE 270 AA; 31491 MW; IDBA6528EADF5 CRC64;
 Query Match	
Best Local Similarity 93.0%; Score 1330.5; DB 2; Length 270;	
Matches 251; Conservative 4; Pred.No. 2,3e-75; Indels 1; Gaps 1;	
OY	1 VRVVPOLQPNPNSQQQPQSEVPLVVOOQFFGGOQQQPPFPQPYRQPFQFBSQQPRLQLQ 60
DB	2 VRVVPOLQPNPNSQQQPQSEVPLVVOOQFFGGOQQPPFPQPYRQPFQFBSQQPRLQLQ 61
OY	61 PPGRLPFPQSPSPPOQPYRQPFQPSQPGQPIISQQAAQQQQQQQQQQQQQQQQQLIQ 119
DB	62 PPGRLPFPQSPPRPQPYRQPFQPSQPGQPIISQQAAQQQQQQQQQQQQQQQQQLIQ 121

AM 1.1
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cheyenne;
 RA Anderson O.D.;
 submitted (MAY-1966) to the EMBL/Genbank/DBJ databases.

RE SEQUENCES FROM AMR.
 RC STRAIN=Cheyenne;
 RA Anderson O.D.;
 RL Submitted (MAR-1996) to the EMBL/genbank/DBJ databases.

DR EMBL U51304; AAA96523.1; -.
DR PIR: S13333; S13333.
DR PIR: T06282; T06282.
DR GO: GO:0045715; F:nutrient reservoir activity; IEA.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001376; Gladin.
DR InterPro: IPR001954; Glia_gluenin.
DR Pfam: PF00234; TRYP_alpha.amy1.1.
DR PRINTS: PR00208; GLIADGUTEN.
DR PRINTS: PR00209; GLIADIN.
DR SMART: SM00499; AAI; 1.
DR SMART: 288 AA; 3325 MW; D7F6B9133283CA2 CRC64; 4.
DR SEQUENCE

Query Match	91.7%;	Score 1312;	DB 2;	Length 288;
Best Local Similarity	92.5%;	Pred. No. 3.2e-72;		
Matches 248;	Conservative 4;	Mismatches 14;	Indels 2;	Gaps 1;

[illegible]

ID	Q41528	PRELIMINARY;	PRT;	267 AA.
AC	Q41528;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
OS	Alpha-gladin.			
DS	Triticum aestivum (Wheat).			
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
CC	Spermatophyta; Magnoliophyta, Liliopsida, Poales; Poaceae; Pooidae;			
CC	Trilicaceae; Trilicium.			
OX	NCBI_taxid=4565;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Cheyenne;			
RA	Anderson O.D.;			
RL	Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; U50984; AAA96276.1; --			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001376; Gladin.			
DR	InterPro; IPR001954; Glia glutenin.			
DR	Pfam; PF00234; Tryp_alpha_amyl1.1.			
DR	PRINTS; PR00208; GIADGUTEN.			
DR	PRINTS; PR00209; GLIADIN.			
DR	SMART; SM00499; AAI; 1			
SC	SEQUENCE 287 AA; 33193 MW; 05F8226749C9B97 CRC64;			

Query Match	90.8%;	Score 1299.5;	DB 2;	Length 287;
Best Local Similarity	92.9%;	Pred. No. 1.8e-71;		
Matches 249; Conservative	3;	Mismatches 13;	Indels 3;	Gaps 3;

Qy 1 VRVPVPLQIQNSQQQPEQEVPLVQQQCFPGQQQCFPPQQPYPQQQPFPSQQPYQLQP 600

	QY	DB
2	VRRPVPQLQPNQBSQGLPEGEVPLVQOQQQF.GGQQQFPFPQOQFPYQPOQ.FSQLEBYLQLQP	79
61	FPPQALPYPQPOSFPQOPYPQOPQYQSGPQOZIS-QQAQAQQQQQQQQQQQQQQQQQQQQ	11.9
80	FPPQOLPYSQOPFRPQOPYPQOPQYQSGPQOZISQQQQQQQQQQQQQQQQQQQQQQQQQQ	13.9
120	QQQLIPCMQVVLQOHNIAPARSQLVQSGTYQLJLHCCQHLMOQIPBQSGCCALHNVVAI	17.9
140	QQQLIPCMQVVLQOHNKHAGRSQVLQSGTYQLRLCCQHLMQIPBQSGCCALHNVVAI	19.9
180	ILHQQK-QQQQPSQVSQQLQOYPYLGQSGFRPSQONPBAQSGVQZQLPQFEETRNL	23.9
200	ILHQQKQQQQQPSQVSQQLQOYPYLGQSGFRPSQONPQTQSGVQZQLPQFEETRNL	25.9
239	ALQTLPMQNVYIAPYCTIAPFGIGTN	266
260	ALQTLPSMCNVYIAPYCTIAPFGIGTN	287

```
Search completed: December 14, 2004, 17:25:26
Job time : 80 secs
```

```

239 ALQTLPMCNVYIAPYCTIAPFGIGTN 266
      |||||:|||||
260 ALQTLPSMCNVYIPYCTIAPFGIGTN 287

```

```
Search completed: December 14, 2004, 17:25:26
Job time : 80 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: December 14, 2004, 17:00:05 ; Search time 76.6667 Seconds
(without alignments)
1244.635 Million cell updates/sec

Title: US-10-089-700-3-W65

Perfect score: 1437

Sequence: 1 VRYVPVQLPQNPSQQQPOE.....CNVYIAPCTIAPFGIFGTN 266

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_23Sep04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003s:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1424	99.1	266	4	AAU01799
2	1424	99.1	266	8	ADH14513
3	1264.5	88.0	290	7	AAE8574
4	1264.5	88.0	290	8	ADP19626
5	464	32.3	369	2	AAW62647
6	450.5	31.4	297	8	AD071669
7	437.5	30.4	298	8	AD071661
8	437.5	30.4	307	7	ADH89338
9	437.5	30.4	307	8	ADG44134
10	285	19.8	1798	4	ABH71695
11	268	18.6	2285	4	ABH63057
12	265	18.4	1162	3	AAV96255
13	265	18.4	1162	3	AAV8500
14	265	18.4	1162	4	AAH62331
15	265	18.4	1162	5	ABH05621
16	265	18.4	1162	8	ADH50506
17	263.5	18.3	757	8	ADH50905
18	263.5	18.3	1069	6	ABO07138
19	263.5	18.3	1069	6	ADJ37233
20	252	17.5	186	7	ADH89336
21	252	17.5	186	7	ADG44132
22	251.5	17.5	260	8	ADG47673
23	249	17.3	905	8	ABG93053
24	249	17.3	905	6	ABH53130
25	249	17.3	905	7	ADK62564

26	246.5	17.2	900	4	ABH62018
27	242.5	16.9	1069	4	ABH61305
28	240	16.7	1013	4	ABH71039
29	238	16.6	358	7	ADH65556
30	235.5	16.4	1142	7	ADH07968
31	231.5	16.1	149	4	AAH72673
32	231	16.1	1761	4	ABH59512
33	230	16.0	160	7	ADH89335
34	230	16.0	160	8	ADG44131
35	230	16.0	4365	6	ABH02252
36	229	15.9	2237	5	ABG70004
37	229	15.9	2703	4	ABH60074
38	227	15.8	158	3	AAV54568
39	227	15.8	2280	4	ABH61650
40	226	15.7	153	3	AAV69495
41	226	15.7	738	5	ABG93140
42	225.5	15.7	1666	7	ADG71666
43	223	15.5	1428	4	ABH70377
44	222.5	15.5	467	8	ADJ76333
45	222.5	15.5	785	8	ADP98983

ALIGNMENTS

RESULT 1
AAU01799
ID AAU01799 standard; protein, 266 AA.
XX
AC AAU01799;
XX
DT 07-SEP-2001 (first entry)
XX
DE Wheat A-gliadin.
XX
KW Wheat; A-gliadin; epitope; coeliac disease; gluten intolerance;
T-cell binding; antagonist; transglutaminase; transgenic plant.
XX
OS Triticum aestivum.
XX
PN WO200125793-A2.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-GB003760.
XX
PR 01-OCT-1999; 99GB-00023306.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Anderson RP, Hill AVS, Jewell DP
XX
DR WPI; 2001-300179/31.
XX
PT Diagnosing coeliac disease or susceptibility to the disease in an individual by detecting in vitro or in vivo T cells which bind immunodominant T cell epitope obtained from naturally occurring homolog of gliadin.
PT
PT Claim 1; Page 52; 107pp; English.
XX
PS The sequence represents wheat A-gliadin. A-gliadin derived peptides of the invention are used to test mammalian (preferably human) susceptibility to coeliac disease (gluten intolerance). The peptides are contacted with a blood sample and T cell recognition measured, a positive T-cell recognition indicating a susceptibility to coeliac disease. The peptides are useful for inducing tolerance in an individual and antagonists to the peptides are useful for treating or preventing coeliac disease in an individual and for producing an antibody specific to them or a wild-type sequence. A mutant gliadin protein (or its fragment of 15 amino acids in length) whose wild-type sequence can be modified by transglutaminase to a sequence that comprises the epitope, but which has been modified in such a way that it does not contain sequence which can

be modified by transglutaminase to a sequence that comprise the epitope
 CC is useful for decreasing the ability of gliadin protein to cause Coeliac
 CC disease. Nucleic acids encoding proteins antagonistic to the T-cell
 CC binding of the epitopes are useful for obtaining a transgenic plant cell
 CC or seed and for the production of a protein. The resultant crop plant is
 CC useful for obtaining a product of a wheat plant, especially grain, which
 CC is optionally processed into flour or another grain product. Food
 CC comprising the antagonistic protein is useful instead of a wild-type
 CC gliadin

Sequence 266 AA:

Query Match 99.1%; Score 1424; DB 4; Length 266;
 Best Local Similarity 99.6%; Pred. No. 8.2e-115;
 Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VRVVPQLQPNPSSQGPQEGVPLVQGGQFPFGQQQFPFPQPPYQPPFPSSQGYLQLQ 60
 1 VRVVPQLQPNPSSQGPQEGVPLVQGGQFPFGQQQFPFPQPPYQPPFPSSQGYLQLQ 60
 61 FPQPLPYPPQPSFPFPQPPYQPPQPSQGPQPSQQAQGGQGGQGGQGGQGGQGGQGGQ 120
 61 FPQPLPYPPQPSFPFPQPPYQPPQPSQGPQPSQQAQGGQGGQGGQGGQGGQGGQGGQ 120
 121 QQLIPCMQDVVLQGNHIAHARSQVLCQSTYQLLQELCCQHLWQIPESQCCAIHNVVAII 180
 121 QQLIPCMQDVVLQGNHIAHARSQVLCQSTYQLLQELCCQHLWQIPESQCCAIHNVVAII 180
 181 LHQGGKQGGQSSQVSPQPLQGYPLGGGSPFSQGNPQAGSVQPPQLPFEEIRNAL 240
 181 LHQGGKQGGQSSQVSPQPLQGYPLGGGSPFSQGNPQAGSVQPPQLPFEEIRNAL 240
 241 QTLPMQCNVYIAPYCTIAPFGIFGTN 266
 241 QTLPMQCNVYIAPYCTIAPFGIFGTN 266

RESULT 2

ADH14513 standard; protein; 266 AA.

ADH14513;
 11-MAR-2004 (first entry)

A-gliadin protein sequence SEQ ID NO:3.

coeliac disease; gliadin; gliadin T cell epitope; gastrointestinal;
 vaccine.

Synthetic.

WO2003104273-A2.

18-DEC-2003.

05-JUN-2003; 2003WO-GB002450.

05-JUN-2002; 2002GB-00012885.

(ISIS-) ISIS INNOVATION LTD.

Anderson RP, Hill AVS, Jewell DP.

WPI; 2004-043640/04.

Preventing or treating coeliac disease comprises administering agent
 PT which are wheat gliadin T cell epitope capable of being recognized by T
 PT cell receptor.

Example 1; SEQ ID NO 3; 177bp; English.

The present invention describes a method (M1) for preventing or treating

coeliac disease. M1 comprises administering an agent (A) comprising a
 CC gliadin T cell epitope, which is capable of being recognised by a T cell
 CC receptor, to an individual. Gliadin is a component of gluten. (A) has
 CC gastrointestinal activity, and can be used in vaccines. The agent (A) can
 CC be used in the preparation of a medicament for treating or preventing
 CC coeliac disease. (A) can also be used in the preparation of a diagnostic
 CC means for use in diagnosing coeliac disease, or susceptibility to coeliac
 CC disease, in an individual, which involves determining whether T cells of
 CC the individual recognise the agent, recognition by the T cells indicating
 CC that the individual has, or is susceptible to, coeliac disease. The
 CC present sequence represents a protein which is used in the
 CC exemplification of the present invention.

Sequence 266 AA:

Query Match 99.1%; Score 1424; DB 8; Length 266;
 Best Local Similarity 99.6%; Pred. No. 8.2e-115;
 Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VRVVPQLQPNPSSQGPQEGVPLVQGGQFPFGQQQFPFPQPPYQPPFPSSQGYLQLQ 60
 1 VRVVPQLQPNPSSQGPQEGVPLVQGGQFPFGQQQFPFPQPPYQPPFPSSQGYLQLQ 60
 61 FPQPLPYPPQPSFPFPQPPYQPPQPSQGPQPSQQAQGGQGGQGGQGGQGGQGGQGGQ 120
 61 FPQPLPYPPQPSFPFPQPPYQPPQPSQGPQPSQQAQGGQGGQGGQGGQGGQGGQGGQ 120
 121 QQLIPCMQDVVLQGNHIAHARSQVLCQSTYQLLQELCCQHLWQIPESQCCAIHNVVAII 180
 121 QQLIPCMQDVVLQGNHIAHARSQVLCQSTYQLLQELCCQHLWQIPESQCCAIHNVVAII 180
 181 LHQGGKQGGQSSQVSPQPLQGYPLGGGSPFSQGNPQAGSVQPPQLPFEEIRNAL 240
 181 LHQGGKQGGQSSQVSPQPLQGYPLGGGSPFSQGNPQAGSVQPPQLPFEEIRNAL 240
 241 QTLPMQCNVYIAPYCTIAPFGIFGTN 266
 241 QTLPMQCNVYIAPYCTIAPFGIFGTN 266

RESULT 3

AAE38574 standard; protein; 290 AA.

AAE38574;

04-DEC-2003 (first entry)

Wheat alpha-2 gliadin protein.

wheat; therapy; coeliac sprue; dermatitis herpetiformis; gluten toxicity;
 KW glutenase; foodstuff; antinflammatory; dermatological; alpha-2 gliadin.

Triticum aestivum.

WO2003068170-A2.

21-AUG-2003.

14-FEB-2003; 2003WO-US004743.

14-FEB-2002; 2002US-0357238P.

14-MAY-2002; 2002US-0380761P.

28-JUN-2002; 2002US-0392782P.

31-OCT-2002; 2002US-0422933P.

20-NOV-2002; 2002US-0428033P.

20-DEC-2002; 2002US-0435881P.

(STRD) UNIV LELAND STANFORD JUNIOR.

Hausch F, Gray G, Shan L, Khosla C;

WPI; 2003-697466/66.

XX	Treating celiac sprue and/or dermatitis herpetiformis comprises
PT	administering to a patient a dose of a glutenase that attenuates gluten
PT	toxicity in the patient.
XX	
PS	Example 2; Fig 4; 69pp; English.
XX	
CC	The present invention relates to a method for treating celiac sprue
CC	and/or dermatitis herpetiformis. The method involves administering to a
CC	patient a dose of a glutenase that attenuates gluten toxicity in the
CC	patient. The method is also useful in treating a foodstuff to render the
CC	foodstuff less toxic to a celiac sprue patient. The present sequence is
CC	wheat alpha-2 gliadin protein used to illustrate the method of the
CC	invention
XX	
SQ	Sequence 290 AA;
XX	
Query Match	88.0%; Score 1264.5; DB 7; Length 290;
Best Local Similarity	84.7%; Pred. No. 5.4e-101;
Matches 243; Conservative	8; Mismatches 15; Indels 21; Gaps 3
QY	1 VRVPFPOQLPQPSQSQPQEQVPLVQOQOFPQOQOQFPQOQFPYPOQPFPSQAPYLQLP 60
Db	2 VRVPFPOQLPQPSQSQPQEQVPLVQOQOFPQOQOQFPQOQFPYPOQPFPSQAPYLQLP 61
QY	61 F-----PQPLPYPOQPSFPPQPPYPOQPOYSQPQOPISSQQAQOQOQO- 105
Db	62 FPOQPLPYPOQPLPYPOQPLPYPOQPFPPQOQPPQOQPOYSQPQOPISSQQAQOQOQO 121
QY	106 ---QQOQOQOQOQILIQILQOQLIPCMQVVLQOHNIHARSQVLQSTYQLQELCCQLMQI 163
Db	122 QKQOQOQOQOQOQILIQILQOQLIPCRDVVLQSHSIATGSSCVLQOQSTYQLVQQLCCQOQMI 181
QY	164 PQSQSCQAIHNVVHAIIILH---QQQKQOQOQPSQVSVFQOQLQOYPLQGGSFRPSQONPQ 219
Db	182 PQSQSCQAIHNVVHAIIILHQQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 241
QY	220 AQGSVQPOQLPQFESIRNLALQTLPMACNVYLAFTCTTAPFIIFGNI 266
Db	242 AQGSVQPOQLPQFESIRNLALQTLPMACNVYLPYCTTAPVGIFFGNI 288
XX	
RESULT 4	
ID	ADP19626
ID	ADP19626 standard; protein; 290 AA.
AC	ADP19626;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Alpha-2-gliadin protein, SEQ ID 28.
XX	
KW	Gluten; Celiac Sprue; wheat; gliadin; alpha-2-gliadin.
XX	
OS	Triticum aestivum.
XX	
PN	WO2004045392-A2.
XX	
PD	03-JUN-2004.
XX	
PF	20-NOV-2003; 2003WO-US037434.
XX	
PR	20-NOV-2002; 2002US-0428033P.
XX	
PA	(STRD) UNIV LEIAND STANFORD JUNIOR.
XX	
PI	Khosla C, Shan L;
XX	
DR	WPI; 2004-460460/43.
XX	
PT	New gluten oligopeptides, useful for diagnosing Celiac Sprue, in
PT	diagnostic assays for detecting antibodies against such oligopeptides, or
PT	for producing antibodies that bind specifically to such oligopeptides.

XX Example 2; Fig 4; 50pp; English.

PS

CC The present invention relates to novel purified gluten oligopeptides. The

CC gluten oligopeptides comprise multiple T cell or B cell epitopes

CC (ADP19608, ADP19616-ADP19623). The gluten oligopeptides are useful in

CC stimulating T cells from Celiac Sprue patients for diagnostic purposes,

CC in diagnostic assays for detecting antibodies against such oligopeptides

CC or for producing antibodies that bind specifically to such oligopeptides

CC The present sequence was used to illustrate the invention.

CC

SC Sequence 290 AA.

QY Query Match 88.0%; Score 1264.5; DB 8; Length 280;

Db Best Local Similarity 84.7%; Pred.No.5,4e-101;

Matches 243; Conservative 8; Mismatches 15; Indels 21; Gaps 3

QY 1 VRYVPQLQPONSQQQPEQGVPLVQGGGFPQQQQQFPQQPPYPQPPFSPSQPYLQLP 60
|||
2 VRVPVPLQPFNSQQQPEQGVPLVQGGGFQQQQPFPQQPPYPQPPFSPSQPYLQLP 61

QY 61 F-----DQPLPYQPSSFPQQPYQPQVYSQPQPISSQQAQQQQQ- 105

Db 62 PPGQPLPYPPQLFYPPQLPYPPQPPFPQPPQPPYQSQPQQPISQQQQQQQQQQ 121

QY 106 --QQQQQQQQQLLQQLILQQQLIFCMVDYLQGNHIAARSQVLQGSYQLQLCGQHLMQI 166
122 QKQQQQQQQQQLLQQLILQQQLIFCRDVYLQHSIAVGSSQVLQGSTYQLWQLCCQQLMQI 181

QY 164 PEGSQCQAHNVVAIIILH----QQKQQQQQSSVSFFQQLQGYFLQGSRFSRSGQNMP 214
182 PEGSRCQAHNVVAIIILHQQQQQQQQQQQQQQLSLVSFGQPPQQYIPSGQSFSPSGQNMP 241

Db 220 AQGSVQPPQLPQFEIRNIALQTLPAMCNVTAPICTTAPFGIGTN 266
242 AQSVCPPQLPQFEIRNIALETLPAMCNVTYPPCTTAPVGIPTGN 288

QY

Db

RESULT 5

AAM62647

ID AAM62647 standard; protein; 369 AA.

AC AAM62647;

DT 17-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 09-OCT-1998 (first entry)

DE Mature durum wheat glutenin protein.

KX Glutenin gene; durum wheat; low-molecular-weight; transgenic durum wheat

KX Triticum turgidum subsp. durum.

OS FR237538-A1.

PN 26-JUN-1998.

PD 18-DEC-1997; 97FR-00016059.

PR 19-DEC-1996; 96IT-MI002663.

PA (ITU)- ITAL MIN UNIV RICERCA SCI & TECNOLOGICA.

PI D Ovidior, Porceddu E, Marchitelli C, Cardelli LE;

DR WPI, 1998-365055/32.

DR N-PsDB; AAV38816.

PT Durum wheat glutenin gene - coding for glutenin protein of low molecular weight.

PS Claim 8; Page 14; 18pp; French.

XX The present sequence represents the mature glutenin protein. The DNA
 CC sequence encoding this protein is isolated from the genomic DNA of
 CC Triticum durum 11. The gene codes for a low-molecular-weight glutenin
 CC protein and can be used to produce transgenic durum wheat plants with
 CC "better quality characteristics" (no details given). (Updated on 25-MAR-
 CC 2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS
 CC field)

XX Sequence 369 AA;

Query Match 32.3%; Score 464; DB 2; Length 369;

Best Local Similarity 40.2%; Pred. No. 9.8e-32; Indels 80; Gaps 14;
 Matches 134; Conservative 32; Mismatches 87;

6 PQLQPNFSCQDP--OEQVPLVQDQPFSCQDQPF--PQDPYPPQPP----- 50
 38 PQQPCSCQDQDPPISQDQDPPFSCQDQDPPFSCQDQDPPFSCQDQDPP 97
 51 ---SQDPLV-----QLQPFQPMIPYQPSFPPQ-----PYPPQPPQYSG 89
 98 PSCQDQDPLVPPQPSFSCQDQDPPFSCQDQDPPFSCQDQDPPFSCQD 157
 90 PQQPISQD 132
 158 QQQPVLPPQPPFSCQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQD 216
 133 QH-----NIAHARSQVLQSTYQLLQELCCGHTMQLPESQCCALHNVAHAIILHQOK 186
 217 QCSFMAMPQSLARSQMTQSSSCHVMQDCCQDLPQIPQDSRYEALIVASITL--OEQ 274
 187 QQQQSSQVSFQQLQGVFLQGSFPPSCQDQDQDQDQDQDQDQDQDQD 234
 275 QQQVSSITQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQD 331
 235 IENLALQTLPAKCNVYIAY--CTTAPFGIFGT 265
 332 MTSIALRTLPKCNMNVPLRYRTTETVPPGV--GT 363

RESULT 6
 ADO71669
 AD071669 standard; protein; 297 AA.
 AC ADO71669;
 12-AUG-2004 (first entry)
 DE Amino acid sequence of a modified glutenin LMW subunit.
 XX low molecular weight subunit; LMW subunit; glutenin;
 KM wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;
 KM gluten intolerance.
 XX Triticum sp.
 OS Synthetic.
 XX EP1424342-A1.
 XX 02-JUN-2004.
 XX 27-NOV-2002; 2002EP-00026461.
 XX 27-NOV-2002; 2002EP-00026461.
 XX (BAKE-) BAKEMARK DEUT GMBH.
 XX (MONS) MONSANTO AGRAR DEUT GMBH.
 XX (UNIF-) UNIFERN GMBH & CO KG.
 XX (PURA-) PURATOS NV.
 XX Hinzmann E, Wieser H, Stahl U;
 XX WPL; 2004-402870/38.

DR N-PSDB; ADO71668.
 XX Novel nucleic acid comprising sequence encoding modified glutenin
 PT polypeptide, useful for preparing modified glutenin polypeptide as
 PT gliadin substitute in foodstuffs such as dough, pastries and wafers.
 XX Claim 16; Fig 11; 43pp; English.

XX The present sequence represents a modified low molecular weight (LMW)
 CC subunit of glutenin. The wild type subunit is designated clone LMW6, and
 CC is isolated from wheat cultivar Cheyenne. The LMW6 polypeptide does not
 CC contain the allergenic epitope QQQP, and shows some minor differences to
 CC published sequences. It therefore represents a new allele for LMW subunit
 CC genes. The LMW6 polypeptide was modified to produce modified glutenin
 CC polypeptides of the invention. In these modified polypeptides one or more
 CC cysteine residues responsible for intermolecular cross-linking through
 CC disulfide bridges are deleted or substituted. The modified glutenin
 CC polypeptide is useful as a gliadin substitute. It is also useful in the
 CC preparation of foodstuffs, such as flour or for the preparation of
 CC pharmaceutical products, such as tablets, where the foodstuffs contain a
 CC considerably reduced amount of gliadin proteins or no gliadin proteins.
 CC pharmaceutical compositions comprising the modified polypeptide of the
 CC invention are useful for treating patients suffering from coeliac disease
 CC or persons who are intolerant to gluten.

XX Sequence 297 AA;

Query Match 31.4%; Score 450.5; DB 8; Length 297;

Best Local Similarity 41.2%; Pred. No. 1.1e-30; Indels 51; Gaps 13;
 Matches 120; Conservative 38; Mismatches 82;

5 VPQLQPNFSCQDP--OEQVPLVQDQPFSCQDQDQDQDQDQDQDQDQDQ 63
 18 IQMERSIGRLRPMQDPLQCKETFP--QQPSSQ--QQQPPQDPPPLQDQDPPFSQ 71
 64 PMLPYPPQPSFPPQPPYQPPQPSQDQDQDQDQDQDQDQDQDQDQD 116
 72 ---QPLFSQKQPPVLPPQPAFSCQDQDQDQDQDQDQDQDQDQDQD 125
 117 QI-----LQQLLIPCMQDVLQCHNIAHARSQVLQSTYQLLQELCCGHTMQLPESQCCQ 170
 126 QLNPKCVPLQDQ---CSPVAMPQH---LARSQMWQSSCNVMQDCCQQLPRIPQSRYE 179
 171 AIHNVAHAIILHQOK-----QQQPSQSV--SFQPLQ--QYPIGQSGF-----RP 213
 180 AIRAIRFSITLQEQDQGFVQPPQDQDQDQDQDQDQDQDQDQDQDQ 239
 214 SQQNPQAGSVQPPQDPPFERRINLALQTLPAKCNVYIAYCTI--APFGI 262
 240 QQQQVQKGFPLQPHQIARLEWMTSIALRTLPKCNMNVPLYSITSAPLGV 290

RESULT 7
 ADO71661
 AD071661 standard; protein; 298 AA.

AC ADO71661;

12-AUG-2004 (first entry)

DE Amino acid sequence of glutenin clone LMW6.

XX low molecular weight subunit; LMW subunit; glutenin;
 KM wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;
 KM gluten intolerance.

XX Triticum sp.

OS Location/Qualifiers

FT Key 25 /note = this residue forms an intermolecular disulfide
 FT Disulfide-bond bond
 FT Disulfide-bond 227

CC No interference between the individual RNA sequences occur. This sequence
CC represents a protein encoded by a target gene used in the method of the
CC invention.

XX Sequence 307 AA;

SC Query Match

30.4%; Score 437.5; DB 7; Length 307;

Best Local Similarity 42.8%; Pred. No. 1.5e-29;

Matches 125; Conservative 36; Mismatches 76; Indels 55; Gaps 16;

CC

13 PSQQQPEQVPLVQQQCP-----GQQQQPFPQPPQPPFPGQGYLQQLPFPQWLP 67

DB

27 PGLERPMQQLPLPQQTFFPQPLFSQQQ-----QLFPQPSFSGQQP-----PFWQQCP 78

CC

68 YPQPSFPQPP-----YPQPPQPSQPPPI-----SQQAQQQQQQQQQQQQQQQI 114

DB

79 FSQQQPLPQPPPSQQQQLVLPQ-QPPFSQQQQLPVLPPQPSPPQQQQQQLVQQQIP 137

CC

115 -LQQLHQQQLIPCMQVVLQGH-NIAH---ARQVYQGSTYQLQLQLCCQHLMQIPRQS 167

DB

138 VVQPSILQQLNPG-KVFLQCCSPVAMPQRLARSGMLQSSCHVMQCCQCCQLPQIPQS 196

CC

168 QCAAHNVVHAIIHQQKQQQQSSQVSPQQLQVPLFGGSRFPSSQNPQ-----219

DB

197 RYEAIRAIIVSYILL-QEQQVQSGISQSQQQPQQ---LGCVSPQQQSQQLGQQPQQ 251

CC

220 ---AGGS-VQPPQLPQFEIRNLALQTLPAMCNVYIAPY-CTIAPFGIFGT 265

DB

252 QQLAQGTFLQPHQIAQLEVMTSIALRLPTMCSVAVPLXRTTTSVPFGV-GT 302

RESULT 9

ADG44134

ADG44134

standard; protein; 307 AA.

ADG44134;

26-FEB-2004 (first entry)

DT

26-FEB-2004 (first entry)

DE

T. aestivum glutenin-1D1 protein.

KW

oil content; plant; storage protein; seed-specific promoter; 2S-albumin;

KW

7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;

KW

oil production; fat production; free fatty acid production; food;

KW

animal feed; pharmaceutical; fine chemical production; glutenin.

OS

Triticum aestivum.

PN

WO200307643-A2.

XX

25-SEP-2003.

XX

17-MAR-2003; 2003WO-EP002733.

XX

20-MAR-2002; 2002DE-01012893.

XX

(BADI) BASF PLANT SCI GMBH.

XX

Bauer J;

XX

WPI; 2004-011485/01.

XX

N-PSDB; ADG44133.

XX

Increasing total oil content of plants, useful e.g. as foods or animal

CC

feeds, by reducing amount of storage proteins, particularly with double-

CC

stranded interfering RNA.

CC

Claim 4; SEQ ID NO 174; 253bp; German.

CC

This invention describes a novel method for increasing the total oil

CC

content of a plant by reducing the amount of at least one storage protein

CC

in the plant (or its tissue, organs, parts or cells) and selecting plants

CC

that have higher total oil content than starting plants. The storage

CC protein is suppressed by introducing antisense RNA, optionally combined

CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding

CC factors directed against storage protein genes, viral sequences that

CC degrade storage protein RNA, constructs that induce homologous

CC recombination of endogenous storage protein genes or mutations into

CC storage protein genes. Most preferably a plant cell is stably transfected

CC with a recombinant expression construct, then regenerated to plants that

CC contain the incorporated sequence. The expression constructs particularly

CC express a seed-specific promoter and they are introduced into plants by

CC standard methods, e.g. via Agrobacterium. The preferred storage proteins

CC of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-

CC prolamines. Transgenic organisms produced by the new method are used for

CC production of oils, fats, free fatty acids or their derivatives, useful

CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence

CC represents a storage protein used to illustrate the method of the

CC invention.

XX Sequence 307 AA;

SC Query Match

30.4%; Score 437.5; DB 8; Length 307;

Best Local Similarity 42.8%; Pred. No. 1.5e-29;

Matches 125; Conservative 36; Mismatches 76; Indels 55; Gaps 16;

CC

13 PSQQQPEQVPLVQQQCP-----GQQQQPFPQPPQPPFPGQGYLQQLPFPQWLP 67

DB

27 PGLERPMQQLPLPQQTFFPQPLFSQQQ-----QLFPQPSFSGQQP-----PFWQQCP 78

CC

68 YPQPSFPQPP-----YPQPPQPSQPPPI-----SQQAQQQQQQQQQQQQQQQI 114

DB

79 FSQQQPLPQPPPSQQQQLVLPQ-QPPFSQQQQLPVLPPQPSPPQQQQQQLVQQQIP 137

CC

115 -LQQLHQQQLIPCMQVVLQGH-NIAH---ARQVYQGSTYQLQLQLCCQHLMQIPRQS 167

DB

138 VVQPSILQQLNPG-KVFLQCCSPVAMPQRLARSGMLQSSCHVMQCCQCCQLPQIPQS 196

CC

168 QCAAHNVVHAIIHQQKQQQQSSQVSPQQLQVPLFGGSRFPSSQNPQ-----219

DB

197 RYEAIRAIIVSYILL-QEQQVQSGISQSQQQPQQ---LGCVSPQQQSQQLGQQPQQ 251

CC

220 ---AGGS-VQPPQLPQFEIRNLALQTLPAMCNVYIAPY-CTIAPFGIFGT 265

DB

252 QQLAQGTFLQPHQIAQLEVMTSIALRLPTMCSVAVPLXRTTTSVPFGV-GT 302

RESULT 10

ABB71695

ID ABB71695 standard; protein; 1798 AA.

XX ABB71695;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polyprotein SEQ ID NO 41877.

XX Drosophila melanogaster developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX

27-SEP-2001.

XX

23-MAR-2001; 2001WO-US009231.

XX

23-MAR-2000; 2000US-0191637P.

XX

PR 11-JUL-2000; 2000US-00614150.

XX

(PEKE) PE CORP NY.

XX

Venter JC, Adams M, Li PWD, Myers EW;

XX

WPI; 2001-656860/75.

DR

```

DR N-PSDB; ABL16798.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PR interactions.
PS
PS Disclosure; SEQ ID NO 41877; 21np + Sequence Listing; English.
CC
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1798 AA;
XX
XX Query Match 19.8%; Score 285; DB 4; Length 1798;
XX Best Local Similarity 38.7%; Pred No 1.7e-15;
XX Matches 92; Conservative 16; Mismatches 82; Indels 48; Gaps 10.
CY 7 QLPQNPSSQQPQEVEPLVQQQGFPGQQGPFPPQFPYRPPQFPPSQQPYLQLQPFPFWL 66
Db 311 CLKQQQQQQQQQ-QQWAPRPQQQQQMAQRPQQQQQQQQPQQQHTRSPKPSF-LQQ----- 361
CY 67 PYRPQSFPFPQFPYRPPQRYSGPQRPISQQCAQQQQQQQQQQQQQQQQQLIQTLQQQLP- 125
Db 362 -----FTTPPLQQQPNQQAQC-IQQQQQQQQQQQQQQQQQQCVLTQQQPPQG 411
CY 126 -CMDVYLQCH-NIAHRSQVLQGSTYQLTGLCCHLMLQIPESQCAIHNVVAIILH 182
Db 412 QQGVYITRHVINTSTAGCQQLIGHMSL-----ALQKQQ-----LLH 450
CY 183 QQQKQQQSSVSFSQQPLQGYPLTGGSFRPSQNPQAQSVQPPQLPQFEIRLAL 240
Db 451 VQQCAQQQPPQQQQQITVVQQLPPAQQQQCLPQHVVQQQ--QPQQV-QFTQQQQIAL 504
XX
XX RESULT 11
XX ABB63057
XX ID ABB63057 standard; protein; 2285 AA.
XX AC ABB63057;
XX AD
XX AE
XX AF
XX AG
XX AH
XX AI
XX AJ
XX AK
XX AL
XX AM
XX AN
XX AO
XX AP
XX AQ
XX AR
XX AS
XX AT
XX AU
XX AV
XX AW
XX AX
XX AY
XX AZ
XX BA
XX BB
XX BC
XX BD
XX BE
XX BF
XX BG
XX BH
XX BI
XX BJ
XX BK
XX BL
XX BM
XX BN
XX BO
XX BP
XX BQ
XX BR
XX BS
XX BT
XX BU
XX BV
XX BW
XX BX
XX BY
XX BZ
XX CA
XX CB
XX CC
XX CD
XX CE
XX CF
XX CG
XX CH
XX CI
XX CJ
XX CK
XX CL
XX CM
XX CN
XX CO
XX CP
XX CQ
XX CR
XX CS
XX CT
XX CU
XX CV
XX CW
XX CX
XX CY
XX CZ
XX DA
XX DB
XX DC
XX DD
XX DE
XX DF
XX DG
XX DH
XX DI
XX DJ
XX DK
XX DL
XX DM
XX DN
XX DO
XX DP
XX DQ
XX DR
XX DS
XX DT
XX DU
XX DV
XX DW
XX DX
XX DY
XX DZ
XX EA
XX EB
XX EC
XX ED
XX EE
XX EF
XX EG
XX EH
XX EI
XX EJ
XX EK
XX EL
XX EM
XX EN
XX EO
XX EP
XX EQ
XX ER
XX ES
XX ET
XX EU
XX EV
XX EW
XX EX
XX EY
XX EZ
XX FA
XX FB
XX FC
XX FD
XX FE
XX FF
XX FG
XX FH
XX FI
XX FJ
XX FK
XX FL
XX FM
XX FN
XX FO
XX FP
XX FQ
XX FR
XX FS
XX FT
XX FU
XX FV
XX FW
XX FX
XX FY
XX FZ
XX GA
XX GB
XX GC
XX GD
XX GE
XX GF
XX GG
XX GH
XX GI
XX GJ
XX GK
XX GL
XX GM
XX GN
XX GO
XX GP
XX GQ
XX GR
XX GS
XX GT
XX GU
XX GV
XX GW
XX GX
XX GY
XX GZ
XX HA
XX HB
XX HC
XX HD
XX HE
XX HF
XX HG
XX HH
XX HI
XX HJ
XX HK
XX HL
XX HM
XX HN
XX HO
XX HP
XX HQ
XX HR
XX HS
XX HT
XX HU
XX HV
XX HW
XX HX
XX HY
XX HZ
XX IA
XX IB
XX IC
XX ID
XX IE
XX IF
XX IG
XX IH
XX II
XX IJ
XX IK
XX IL
XX IM
XX IN
XX IO
XX IP
XX IQ
XX IR
XX IS
XX IT
XX IU
XX IV
XX IW
XX IX
XX IY
XX IZ
XX JA
XX JB
XX JC
XX JD
XX JE
XX JF
XX JG
XX JH
XX JI
XX JJ
XX JK
XX JL
XX JM
XX JN
XX JO
XX JP
XX JQ
XX JR
XX JS
XX JT
XX JU
XX JV
XX JW
XX JX
XX JY
XX JZ
XX KA
XX KB
XX KC
XX KD
XX KE
XX KF
XX KG
XX KH
XX KI
XX KJ
XX KK
XX KL
XX KM
XX KN
XX KO
XX KP
XX KQ
XX KR
XX KS
XX KT
XX KU
XX KV
XX KW
XX KY
XX KZ
XX LA
XX LB
XX LC
XX LD
XX LE
XX LF
XX LG
XX LH
XX LI
XX LJ
XX LK
XX LL
XX LM
XX LN
XX LO
XX LP
XX LQ
XX LR
XX LS
XX LT
XX LU
XX LV
XX LW
XX LX
XX LY
XX LZ
XX MA
XX MB
XX MC
XX MD
XX ME
XX MF
XX MG
XX MH
XX MI
XX MJ
XX MK
XX ML
XX MM
XX MN
XX MO
XX MP
XX MQ
XX MR
XX MS
XX MT
XX MU
XX MV
XX MW
XX MX
XX MY
XX MZ
XX NA
XX NB
XX NC
XX ND
XX NE
XX NF
XX NG
XX NH
XX NI
XX NJ
XX NK
XX NL
XX NM
XX NO
XX NP
XX NQ
XX NR
XX NS
XX NT
XX NU
XX NV
XX NW
XX NX
XX NY
XX NZ
XX OA
XX OB
XX OC
XX OD
XX OE
XX OF
XX OG
XX OH
XX OI
XX OJ
XX OK
XX OL
XX OM
XX ON
XX OO
XX OP
XX OQ
XX OR
XX OS
XX OT
XX OU
XX OV
XX OW
XX OX
XX OY
XX OZ
XX PA
XX PB
XX PC
XX PD
XX PE
XX PF
XX PG
XX PH
XX PI
XX PJ
XX PK
XX PL
XX PM
XX PN
XX PO
XX PP
XX PQ
XX PR
XX PS
XX PT
XX PU
XX PV
XX PW
XX PX
XX PY
XX PZ
XX QA
XX QB
XX QC
XX QD
XX QE
XX QF
XX QG
XX QH
XX QI
XX QJ
XX QK
XX QL
XX QM
XX QN
XX QO
XX QP
XX QQ
XX QR
XX QS
XX QT
XX QU
XX QV
XX QW
XX QX
XX QY
XX QZ
XX RA
XX RB
XX RC
XX RD
XX RE
XX RF
XX RG
XX RH
XX RI
XX RJ
XX RK
XX RL
XX RM
XX RN
XX RO
XX RP
XX RQ
XX RR
XX RS
XX RT
XX RU
XX RV
XX RW
XX RX
XX RY
XX RZ
XX SA
XX SB
XX SC
XX SD
XX SE
XX SF
XX SG
XX SH
XX SI
XX SJ
XX SK
XX SL
XX SM
XX SN
XX SO
XX SP
XX SQ
XX SR
XX SS
XX ST
XX SU
XX SV
XX SW
XX SX
XX SY
XX SZ
XX TA
XX TB
XX TC
XX TD
XX TE
XX TF
XX TG
XX TH
XX TI
XX TJ
XX TK
XX TL
XX TM
XX TN
XX TO
XX TP
XX TQ
XX TR
XX TS
XX TT
XX TU
XX TV
XX TV
XX TW
XX TX
XX TY
XX TZ
XX UA
XX UB
XX UC
XX UD
XX UE
XX UF
XX UG
XX UH
XX UI
XX UJ
XX UK
XX UL
XX UM
XX UN
XX UO
XX UP
XX UQ
XX UR
XX US
XX UT
XX UU
XX UV
XX UV
XX UW
XX UX
XX UY
XX UZ
XX VA
XX VB
XX VC
XX VD
XX VE
XX VF
XX VG
XX VH
XX VI
XX VJ
XX VK
XX VL
XX VM
XX VN
XX VO
XX VP
XX VQ
XX VR
XX VS
XX VT
XX VU
XX VW
XX VX
XX VY
XX VZ
XX WA
XX WB
XX WC
XX WD
XX WE
XX WF
XX WG
XX WH
XX WI
XX WJ
XX WK
XX WL
XX WM
XX WN
XX WO
XX WP
XX WQ
XX WR
XX WS
XX WT
XX WU
XX WV
XX WX
XX WY
XX WZ
XX XA
XX XB
XX XC
XX XD
XX Xe
XX XF
XX XG
XX XH
XX XI
XX XII
XX XIII
XX XIV
XX XV
XX XVI
XX XVII
XX XVIII
XX XIX
XX XX
XX XXI
XX XXII
XX XXIII
XX XXIV
XX XXV
XX XXVI
XX XXVII
XX XXVIII
XX XXIX
XXX

```

PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
PS	Disclosure; SEQ ID NO 15963; 21bp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB10511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB85737-AB85720). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPD at ftp.wipd.int/pub/published_pct_sequences
CC	
CC	
CC	
CC	
SQ	Sequence 2285 AA:
Query Match	18.6%; Score 268; DB 4; Length 2285;
Best Local Similarity	35.7%; Pred. No. 6.6e-14;
Matches 101; Conservative % 14; Mismatches 104; Indels 64; Gaps 13;	
Oy	7 OLGRNP-----SQRRGEVPLVQQQRPRGQQQRPFRPQPQ---PP-SQQP 54
Db	909 QMGGQGPVAVTNVHGVMPQQQVA--QQQQQPMQMQLPVGVYQVPFTVLPRPHSDQP 966
Oy	55 YLGLRPFQ-----FWLPYPQPGSFPQRPQRPQRPQYSGPDPISSQQA--- 99
Db	967 QQQQQPLQQQLQGMLHTNVGAFL--TQQQMAQQQAQGYFQQQQQQCGQAVMNQAAVAM 1024
Oy	100 CQQQQCQQQQCQQCQLLQQLLQQLI-----PCMDVLLQGNIAHARQYLQGSYYQLQ 153
Db	1025 QQAQQQQQLSLSPDLIQQQILQQQQGVAVSHQQQIMQQQLAQHQQLQQLQQQQQLQQQQ 1084
Oy	154 ELCCOHLMQPREQSQCALHNHVAILHQQCKQQQQQSSQV----SFQCELQ-----QY 204
Db	1085 QIQQQQLQQLQQLQQQ-QVQCYAQA--MPQQCHQQLVTGSQVMAFHCHQQLQLPVMQNY 1144
Oy	205 PL-----GQGSFRPSQCNPAQGSVQPPQLDPQ 232
Db	1142 PRTSVAPRIQHTVNGGGQVTLSDAQCCQHDFGFAVPPQQAAPF 1184
RESULT 12	
AAAY96255	
ID	AAAY96255 standard; protein; 1162 AA.
AC	XX
XX	XX
DT	12-SEP-2003 (revised)
DT	11-SEP-2000 (first entry)
XX	XX
DE	Kaposi's sarcoma-associated herpesvirus LANA.
XX	XX
KW	Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus; latency-associated nuclear antigen; LANA; gamma-2 herpes virus; Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVC4E; Kaposi's sarcoma; primary effusion lymphoma; PEL; human immunodeficiency virus; HIV; multicentric Castlemann's disease.
KX	XX
OS	Human herpesvirus 8.
XX	XX
FH	Key Location/Qualifiers
FT	Domain 14..17 /note= "nuclear localisation signal, NLS"
FT	Domain 64..70 /note= "nuclear localisation signal, NLS"
FT	Region 320..429 /note= "acidic repeat region"
FT	Region 430..549 /note= "Gln, Glu, Pro-rich region"
FT	Region 550..589

N
U
C

— 7
X
7

:: b
:: d

N
U
d

Db 495 PLQEPQQEFPQQEFPQQEFPQQEFPQQEFPQQEFPQQE-PPQQEFPQRE 55

Sequence 1162 AA;

The present invention describes a system (A) for maintaining a plasmid as an episome in mammalian cells, comprising the rhadoviral sequence LMAA (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487, S1) expressed in the cell, and the rhadoviral sequence RVCA8 (rhadoviral cis-acting element) of 801 base pairs (see ABA93488, S2) present in the plasmid. Also describes is a method for maintaining a closed circular DNA in a cell by expressing (S1) in the cells and having (S2) as a cis-acting and maintenance sequence in the DNA. (A) is particularly used in gene therapy (or other gene transfer applications) that uses mammalian cells in which LMAA is expressed. (A) improves persistence of gene therapy vectors in cells. The present sequence

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 Seconds

(without alignments)
928.452 Million cell updates/sec

Title: US-10-089-700-3-w65

Perfect score: 1437

Sequence: 1 VRRVVPQLQPNPSQQQPE.....CNVYIAFYCTIAFPGIFGTN 266

Scoring table: BLOSUM62

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/SA_COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/SA_COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/SA_COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/SA_COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/SA_COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	464	32.3	369	2	US-08-991-300-2
2	265	18.4	1162	2	US-08-728-323A-2
3	265	18.4	1162	3	US-09-298-568-2
4	265	18.4	1162	4	US-09-410-399-2
5	265	18.4	1162	4	US-09-894-273-2
6	247	17.2	788	2	US-08-918-914-4
7	236	16.4	498	4	US-09-270-767-45042
8	234	16.3	256	4	US-09-248-796A-21251
9	215.5	15.0	579	4	US-09-668-119-3
10	215	15.0	2074	4	US-09-491-356C-9
11	201	14.0	2023	4	US-09-491-356C-8
12	201	14.0	2124	4	US-09-538-092-1377
13	198.5	13.8	505	4	US-09-248-796A-19253
14	197.5	13.7	663	4	US-09-270-767-61220
15	197.5	13.7	1591	4	US-09-270-767-45698
16	197.5	13.7	2441	1	US-08-194-468-2
17	197.5	13.7	2441	1	US-08-961-739-2
18	197.5	13.7	2441	3	US-09-514-247A-8
19	197.5	13.7	2441	3	US-09-686-316-2
20	196.5	13.7	379	4	US-09-248-796A-23759
21	196.5	13.7	2442	4	US-09-514-247A-10
22	196.5	13.7	2442	4	US-09-538-092-1370
23	196	13.6	216	4	US-09-248-796A-21017
24	194.5	13.5	729	3	US-09-625-188-20
25	185	12.9	295	4	US-09-248-796A-20004
26	184.5	12.8	316	4	US-09-270-767-42663
27	184.5	12.8	320	4	US-09-248-796A-24758

28	184.5	12.8	332	4	US-09-248-796A-21649	Sequence 21649, A
29	183	12.7	169	4	US-09-248-796A-28087	Sequence 28087, A
30	183	12.7	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	1319	4	US-09-538-092-1291	Sequence 1291, Ap
34	177	12.3	261	4	US-09-602-565-34	Sequence 34, Appl
35	176.5	12.3	673	4	US-09-248-796A-20659	Sequence 20659, A
36	176.5	12.3	684	4	US-09-823-240A-9	Sequence 9, Appli
37	176	12.2	618	4	US-09-248-796A-15319	Sequence 15319, A
38	176	12.2	848	4	US-09-538-092-33	Sequence 33, Appl
39	174.5	12.1	382	4	US-09-248-796A-18720	Sequence 18720, A
40	174	12.1	657	4	US-09-248-796A-19232	Sequence 19232, A
41	173	12.0	1507	4	US-09-914-359-37	Sequence 37, Appl
42	171.5	11.9	903	2	US-08-853-310-2	Sequence 2, Appli
43	168.5	11.7	311	4	US-09-248-796A-27827	Sequence 27827, A
44	167.5	11.7	667	2	US-08-718-661-2	Sequence 2, Appli
45	167	11.6	542	1	US-07-814-564-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-991-300-2
Sequence 2, Application US/08991300
Patent No. 5973225
GENERAL INFORMATION:
APPLICANT: D/OVIDIO, RENATO
APPLICANT: PORCEDDU, ENRICO
APPLICANT: MERCHITELLI, CINZIA
APPLICANT: CARDELLI, LUISA ERCOLI
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE
TITLE OF INVENTION: ENCODING A LOW MOLECULAR WEIGHT GLUTENIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: OBOLON, SPYVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,300
FILING DATE: 16-DEC-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT MI 96/A 002663
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBOLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2264-0201-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-300-2
Query Match 32.3%; Score 464; DB 2; Length 369;
Best Local Similarity 40.2%; Pred. No. 2.3e-35;

US-09-410-399-2
Sequence 2, Application US/09410399
Patent No. 6482587
GENERAL INFORMATION:
APPLICANT: Robertson, Erle S.
APPLICANT: Corter, Murray A.
TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
FILE REFERENCE: UM-03778
CURRENT APPLICATION NUMBER: US/09/410,399
CURRENT FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-2

Query Match 18.4%; Score 265; DB 4; Length 1162;
Best Local Similarity 38.0%; Pred. No. 3, 1e-16;
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

QY 4 PVPOLQPONPSQQQPPQEVPLVQ--QQQFPQQ--QQFPQQPPYPPQPPFPSCQPYLQ 59
DB 495 PLOGEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEP 553
QY 60 P---FPQ---PMLPYPPQ---PQSFPPQPPYPPQ--QPQYSQPQQPPISSQQAQQQQQQQQQQ 109
DB 554 PQGREPQQRPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREP 613
QY 110 QQQQILQQILQQILPCMDVVLQOHNIHARSQVLTQSTYQLJQLCCQHLWQIPQSGQC 169
DB 614 EQQDDEQQDDEQ---QDEQQDDEQQDDEQQDDEQQDDEQQDDE---QQDDEQQCD 663
QY 170 QAIHNVHAILLHQQKQKQKQSSQVSFQQPLQYPLGGGSRPPSQNPAQGSVQPPQL 229
DB 664 EQQCD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
QY 230 PQFEE 234
DB 717 EQQCD 721

RESULT 5
US-09-894-273-2
Sequence 2, Application US/09894273
Patent No. 6756203
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballesetas, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS IANA ACTS IN TRANS ON A UNIT OF RHADINO
TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EROSOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/894,273
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/109,422
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2

Query Match 18.4%; Score 265; DB 4; Length 1162;
Best Local Similarity 38.0%; Pred. No. 3, 1e-16;
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;
QY 4 PVPOLQPONPSQQQPPQEVPLVQ--QQQFPQQ--QQFPQQPPYPPQPPFPSCQPYLQ 59
DB 495 PLOGEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEP 553
QY 60 P---FPQ---PMLPYPPQ---PQSFPPQPPYPPQ--QPQYSQPQQPPISSQQAQQQQQQQQQQ 109
DB 554 PQGREPQQRPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREP 613
QY 110 QQQQILQQILQQILPCMDVVLQOHNIHARSQVLTQSTYQLJQLCCQHLWQIPQSGQC 169
DB 614 EQQDDEQQDDEQ---QDEQQDDEQQDDEQQDDEQQDDEQQDDE---QQDDEQQCD 663
QY 170 QAIHNVHAILLHQQKQKQKQSSQVSFQQPLQYPLGGGSRPPSQNPAQGSVQPPQL 229
DB 664 EQQCD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
QY 230 PQFEE 234
DB 717 EQQCD 721

DB 495 PLOGEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEP 553
QY 60 P---FPQ---PMLPYPPQ---PQSFPPQPPYPPQ--QPQYSQPQQPPISSQQAQQQQQQQQQQ 109
DB 554 PQGREPQQRPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREPQQREP 613
QY 110 QQQQILQQILQQILPCMDVVLQOHNIHARSQVLTQSTYQLJQLCCQHLWQIPQSGQC 169
DB 614 EQQDDEQQDDEQ---QDEQQDDEQQDDEQQDDEQQDDEQQDDE---QQDDEQQCD 663
QY 170 QAIHNVHAILLHQQKQKQKQSSQVSFQQPLQYPLGGGSRPPSQNPAQGSVQPPQL 229
DB 664 EQQCD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
QY 230 PQFEE 234
DB 717 EQQCD 721

RESULT 6
US-08-918-914-4
Sequence 4, Application US/08918914
Patent No. 6876963
GENERAL INFORMATION:
APPLICANT: Mitchell, Peter
APPLICANT: Hutchinson, Nancy
APPLICANT: Lawton, Michael
APPLICANT: Magna, Holly
APPLICANT: Yocum, Sue
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,914
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy U.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1070094
US-08-918-914-4

Query Match 17.2%; Score 247; DB 2; Length 788;
Best Local Similarity 31.7%; Pred. No. 9, 1e-15;
Matches 91; Conservative 21; Mismatches 97; Indels 76; Gaps 11;

RESULT 14
 US-09-270-767-61220
 Sequence 61220, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 FILE REFERENCE: File Reference: 7325-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517

Query Match	13.7%	Score 197.5	DB 4	Length 1591
Best Local Similarity	29.8%	Pred. No. 8.6%		
Matches	88	Conservative	25	Mismatches 97; Indels 85; Gaps 13;

QY	4	PVPOLQPNESQCPQPCQCVPLVYQQQCTPFGQQQQPFPQQPQYQPPQFPFSQQPYQLQQLPPQ	63
DB	1123	PVAE--EEQFQDEAQRVAVSTQPPQQLQ-----QQPFRQYQGPY-QL-PLP-	1170
QY	64	PMLPYPQPSPPQCPYPQPPQYQSQCPQPSQQ-----AQQQQQQ	105
DB	1171	--LPAQRHSNPP-----QQQQQHQPQYVPEEQFLKILEEHLQARAYHQLQQQQQH	1224
QY	106	QQQQQQQQQLLQQLIQQLIPGMDVVLQCHNIAHRSVQLQSTQQL-----LQE	154
DB	1225	QQQQQQQQQHQQQQQQQH-----HKQLPIHSTA-ATKAVLQQAQDSLGLGYSRERFVAQE	1278
QY	155	LC-----CQHLMQIPESQCAIHNVVAIILHQQKQKQQQPSQVS	196
DB	1279	LVTPTVSHPRGGPKYLLPQKQIQCEBDEQQQQQPCFQVQLHKPDPHQEQQLIHGLPPIA	1338
QY	197	EQQLQYPLPGQSFRPQSQQPKAQSGVQPCQLPQFEIRMLAQUTLPAMCNVYI	251
DB	1339	YYQP-----QISYKTLPNHPLAKSSLE-----SEIEKLIAANKPGSLAV	1379

Wed Dec 15 10:02:02 2004

us-10-089-700-3-w65.ra1

Page 7

Search completed: December 14, 2004, 17:29:02
Job time : 20 secs

This Page Blank (uspto)

APPLICANT: Drifhout, Jan W.
 APPLICANT: Koning, Frits
 APPLICANT: Mcadam, Stephan N.
 APPLICANT: Ludwig, Solid Magne
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
 TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
 FILE REFERENCE: 2799/71244-PCT-US
 CURRENT APPLICATION NUMBER: US/10/474,955
 CURRENT FILING DATE: 2003-10-13
 NUMBER OF SEQ ID NOS: 137
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 99
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Amino acid sequence of GAMMA-4
 US-10-474-955-99

Query Match 36.3%; Score 521; DB 17; Length 279;
 Best Local Similarity 46.2%; Pred. No. 3,4e-33;
 Matches 132; Conservative 26; Mismatches 76; Indels 52; Gaps 14;

4 PVPQLQPNPSSQCPQEQVPLVQCCQPPGQQQQCP--PQCPYPQPPFPSPGQPYLQLPF 61
 18 PVP--QPHQPSQCP-----QCTFPQQTFFHQPPQCPQCP--PQCPFLQPPQPF 66
 62 P-QPWLPPQPSFPQPPQPPQ-----PQYSPQCPISQQAQ-----QQQQQQQQQQ 110
 67 PQCPQPPYPQ-----QPQPPFPQTPQQLFPQSQCPQCPQCPQCPQCPQCPQCPQCP 122
 111 QQQI 166
 123 QPFPQPSLQCCQVNCNFKFLQCCCKPVSLSMSWMPQSDQVMRQSCQQLAQIPQ 182
 167 SCCAIHNVVAIILHQCKKQCCQPSQVSPQPL-QQYPLGGGFRPSSQNPQAQGSVQ 225
 183 LCCAIHTVHSIIMQEQEQ-----GMHILLPLVQQQGVQGGTL-----VQGGGIIP 230
 226 PQCLPQFEIRNLALQTLPMCNVYIAPYCTI--APF-----GIFG 264
 231 PQCPALQELNLSLVLTPLTMCNVVPECCITKAPSSVVAIGG 276

RESULT 9
 US-10-474-955-98
 Sequence 98, Application US/10474955
 Publication No. US20040241161A1
 GENERAL INFORMATION:
 APPLICANT: Drifhout, Jan W.
 APPLICANT: Koning, Frits
 APPLICANT: Mcadam, Stephan N.
 APPLICANT: Ludwig, Solid Magne
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
 TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
 FILE REFERENCE: 2799/71244-PCT-US
 CURRENT APPLICATION NUMBER: US/10/474,955
 CURRENT FILING DATE: 2003-10-13
 NUMBER OF SEQ ID NOS: 137
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 98
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Amino acid sequence of GAMMA-2
 US-10-474-955-98

Query Match 36.0%; Score 518; DB 17; Length 279;
 Best Local Similarity 45.8%; Pred. No. 5,8e-33;
 Matches 131; Conservative 26; Mismatches 77; Indels 52; Gaps 14;

4 PVPQLQPNPSSQCPQEQVPLVQCCQPPGQQQQCP--PQCPYPQPPFPSPGQPYLQLPF 61

Db 18 PVP--QPHQPSQCP-----QCTFPQQTFFHQPPQCPQCP--PQCPFLQPPQPF 66
 QY 62 P-QPWLPPQPSFPQPPQPPQ-----PQYSPQCPISQQAQ-----QQQQQQQQQQ 110
 Db 67 PQCPQPPYPQ-----QPQPPFPQTPQQLFPQSQCPQCPQCPQCPQCPQCPQCPQCP 122
 QY 111 QQQI 166
 Db 123 QPFPQPSLQCCQVNCNFKFLQCCCKPVSLSMSWMPQSDQVMRQSCQQLAQIPQ 182
 QY 167 SCCAIHNVVAIILHQCKKQCCQPSQVSPQPL-QQYPLGGGFRPSSQNPQAQGSVQ 225
 Db 183 LCCAIHTVHSIIMQEQEQ-----GMHILLPLVQQQGVQGGTL-----VQGGGIIP 230
 QY 226 PQCLPQFEIRNLALQTLPMCNVYIAPYCTI--APF-----GIFG 264
 Db 231 PQCPALQELNLSLVLTPLTMCNVVPECCITKAPSSVVAIGG 276

RESULT 10
 US-10-474-955-100
 Sequence 100, Application US/10474955
 Publication No. US20040241161A1
 GENERAL INFORMATION:
 APPLICANT: Drifhout, Jan W.
 APPLICANT: Koning, Frits
 APPLICANT: Mcadam, Stephan N.
 APPLICANT: Ludwig, Solid Magne
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
 TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
 FILE REFERENCE: 2799/71244-PCT-US
 CURRENT APPLICATION NUMBER: US/10/474,955
 CURRENT FILING DATE: 2003-10-13
 NUMBER OF SEQ ID NOS: 137
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 100
 LENGTH: 279
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Amino acid sequence of GAMMA-3
 US-10-474-955-100

Query Match 36.0%; Score 518; DB 17; Length 279;
 Best Local Similarity 46.0%; Pred. No. 5,8e-33;
 Matches 131; Conservative 28; Mismatches 76; Indels 50; Gaps 14;

QY 4 PVPQLQPNPSSQCPQEQVPLVQCCQPPGQQQQCP--PQCPYPQPPFPSPGQPYLQLPF 61
 Db 18 PVP--QPHQPSQCP-----QCTFPQQTFFHQPPQCPQCP--PQCPFLQPPQPF 66
 QY 62 P-QPWLPPQ-----PQYSPQCPISQQAQ-----QQQQQQQQQQ 111
 Db 67 PQCPQPPYPQ-----QPQPPFPQTPQQLFPQSQCPQCPQCPQCPQCPQCPQCPQCP 123
 QY 112 QQQI 167
 Db 124 PFPQPSLQCCQVNCNFKFLQCCCKPVSLSMSWMPQSDQVMRQSCQQLAQIPQ 183
 QY 168 SCCAIHNVVAIILHQCKKQCCQPSQVSPQPL-QQYPLGGGFRPSSQNPQAQGSVQ 226
 Db 184 LCCAIHTVHSIIMQEQEQ-----GMHILLPLVQQQGVQGGTL-----VQGGGIIP 231
 QY 227 PQCLPQFEIRNLALQTLPMCNVYIAPYCTI--APF-----GIFG 264
 Db 232 PQCPALQELNLSLVLTPLTMCNVVPECCITKAPSSVVAIGG 276

RESULT 11
 US-10-739-930-9778
 Sequence 9778, Application US/10739930
 Publication No. US20040216190A1

SEQ ID NO 200100
LENGTH: 541
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(541)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: WRT4577_114080C.1.dep
US-10-425-115-200100

Query Match 19.3%; Score 277; DB 17; Length 541;
Best Local Similarity 37.5%; Pred. No. 1,2e-13;
Matches 90; Conservative 20; Mismatches 80; Indels 50; Gaps 11;

QY 9 QPONSQQQPOBOVFLVQQQQPPGQQQPPQPPYPPQPPPSQQPYLQLQPPFPWLPY 68
DB 328 KPQASTQUTPMQ---QQLQPFQQQQQL-QQHNMPPQSLPLQSSQMLQ-----Q 375
QY 69 PGPQSFPPQPPY---QPQYSSQP-QPISQQQAQQQQQQQQQQQQQQQLLQQLL 124
DB 376 QQPQMQPQQQQPQQ 433
QY 125 PCMDVVLQGHNIHARSGVLAQSTYQLLDELCCQHLWQIPESQSCQAHHNVHAHLLHQ 184
DB 434 -----QQQQQQQLPQQQQQQQQQQQQQQQQQQ-----QPQQQQQQM-----QQ 470
QY 185 QKQQQQPPSSQVSPFQQPLQCYPLGGGFRP-SQNPQAQGSVQPQQLPQFBEIRNLALQT 243
DB 471 QQQQQMQPQ-----QQQQQQQQQQQQQQQQQQQQPQQQQQPPQVGTGMGQ---QFMQGHNRRAVQM 522

RESULT 15

US-10-425-115-200097
Sequence 200097, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5322)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 200097

LENGTH: 283

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(283)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRP4577_114078C.1.dep

US-10-425-115-200097

Query Match 19.1%; Score 274; DB 17; Length 283;
Best Local Similarity 38.2%; Pred. No. 1.1e-13;
Matches 92; Conservative 20; Mismatches 71; Indels 58; Gaps 13;

QY 9 QPONSQQQPOBOVFLVQQQQPPGQQQPPQPPYPPQPPPSQQPYLQLQPPFPWLPY 68
DB 76 KPQASTQUTPMQ---QQLQPFQQQQQL-QQHNMPPQSLPLQSSQMLQ-----Q 123
QY 69 PGPQSFPPQPPY---QPQYSSQP-QPISQQ-QAQQQQQQQQQQQQQQQLLQQLL 123
DB 124 QQPQMQPQQQQPQQ 183
QY 124 PCMDVVLQGHNIHARSGVLAQSTYQLLDELCCQHLWQIPESQSCQAHHNVHAHLLHQ 183

DB 184 QP-----QQQQM-----QQQQQQQQQQM-----PQQQQM-----Q 211
QY 184 QKQQQQPPSSQVSPFQQPLQCYPLGGGFRP-SQNPQAQGSVQPQQLPQFBEIRNLALQT 242
DB 212 QQQQQMQPQ-----QQQQQQQQQQQQQQQQQQQQPQQQQQPPQVGTGMGQ---QFMQGHNRRAVQM 263
QY 243 L 243
DB 264 M 264

Search completed: December 14, 2004, 17:35:36
Job time : 65.8333 secs

RESULT 2
EMBL
alpha/beta-gliadin precursor - wheat
N/Alternate names: prolamin
C/Species: Triticum aestivum (common wheat)
C/Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004
C/Accession: A03354
Rafalski, J.A.; Scheefs, K.; Metzler, M.; Peterson, D.M.; Hedgcock, C.; Soll, D.G.
EMBL J. 3, 1409-1415, 1984
/Title: Developmentally regulated plant genes: the nucleotide sequence of a wheat gliadin

RESULT 12

A:Residues: 1-292 <OKI>
A:Cross-references: UNIPROT:P04721
C:Superfamily: gliadin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-292/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 80.4%; Score 1155.5; DB 2; Length 292;
Best Local Similarity 84.1%; Pred. No. 1,1e-70;
Matches 227; Conservative 3; Mismatches 9; Indels 31; Gaps 4;

```
QV 1 VRVPVPLQPNPSSQOPQEQVPLVQQQQFPQQQPPQPPYQPPQPPSSQQPYLQLLP 60
DB 21 VRVPVPLQPNPSSQOPQEQVPLVQQQQFPQQQPPQPPYQPPQPPSSQQPYLQLLP 80
QV 61 F--PQPV---LPYQPSFPQPPQPPQPPQPPIS-QQQAQQQQQQQQQQQQQQ 114
DB 81 FLQPPFPFPQLPYSQPPFPQPPQPPQPPQPPISQQQQQQQQQQQQQQQQQQ 140
QV 115 LQQILQQQLPCMDVVLQGHNTAHARSQVLQGSTYQLLQELCCQHLWQIPQSSCCQAIHN 174
DB 141 IQQILQQQLPCMDVVLQGHNTAHARSQVLQGSTYQLLQELCCQHLWQIPQSSCCQAIHN 200
QV 175 VVHAIIILH-----QQQRQQQPPSSQVSFPQLQYPLGGG 209
DB 201 VVHAIIILHQQQQQQEQKQLQQQQQQQQQQQQQQQQKQQQQPPSSQVSFPQLQYPLGGG 260
QV 210 SFRPSQNPQAQGSVPQQLPQPEIRNTA 239
DB 261 SFRPSQNPQAQGSVPQQLPQPEIRNTA 290
```

Search completed: December 14, 2004, 17:27:00
Job time : 16 secs

FT	VAR	SEQUENCE	286 AA;	193	194	MM;	ESECPABBE29E10C6	CR64;	HN -> LK (in Ref. 3)
SC	SEQUENCE	286 AA;	132949	MM;	ESECPABBE29E10C6	CR64;			
Query Match	Best Local Similarity	94.2%;	Score 1353;	DB 1;	Length 286;				
Matches	254;	Conservative	0;	Mismatches	12;	Indels	0;	Gaps	0;
QY	1	VRVVPVLQQLPQNFQSQQFQREAVPLVQQQCFEPQDQDQFPPOQFYPQDPFFSQQPYLQLQP	60						
DB	21	VRFPVPLQQLPQNFQSQQLPQEQVPLVQQQQLFGQQQFPFPQDPYPQDPFFSQQLPYLQLQP	80						
QY	61	FPQFWLFPQDQSFPPQDPFYPQDPQYSQDQPISSQDQAQQQQQQQQQQQQQQQLLQQLLQ	120						
DB	81	FPQQLPYSQDPFRPQDPYPQDPQYSQDQPISSQDQAQQQQQQQQQQQQQQQLLQQLLQ	140						
QY	121	QQLPQCMQDVVLQENIHARSGQVLCQSTYQLLQELCCQHLMOIPESQCCQAIHNVYHAI	180						
DB	141	QQLPQCMQDVVLQENIHARSGQVLCQSTYQLLQELCCQHLMOIPESQCCQAIHNVYHAI	206						
QY	181	LHQQKQKQKQSSQVSFQDPLQDPLQDQGSFRPQDQNPQAQGSVQDQQLQFEEIRNAL	240						
DB	201	LHQQKQKQKQSSQVSFQDPLQDQDPLQDQGSFRPQDQNPQAQGSVQDQQLQFEEIRNAL	260						
QY	241	QTLPAMCNVYIAPYCTIAPFGIFGTN	266						
DB	261	QTLPAMCNVYIAPYCTIAPFGIFGTN	286						
RESULT 3									
AAA96525									
ID	AAA96525	PRELIMINARY;	PRT;	286	AA.				
AC	AAA96525;								
DT	02-MAR-2004	(TREMBlrel. 27, Created)							
DT	02-MAR-2004	(TREMBlrel. 27, Last sequence update)							
DT	02-MAR-2004	(TREMBlrel. 27, Last annotation update)							
DE	Alpha-glucan storage protein.								
OS	Triticum aestivum (wheat).								
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;								
OX	Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.								
NCBI	_taxid=4565;								
RM	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Cheyenne;								
RA	Anderson O.D.;								
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.								
DR	EMBL, U51307; AAA96525.1; -								
SD	SEQUENCE 286 AA; 32949 MM; BSECPABBE29E10C6	CRC64;							
Query Match	Best Local Similarity	94.2%;	Score 1353;	DB 2;	Length 286;				
Matches	254;	Conservative	0;	Mismatches	12;	Indels	0;	Gaps	0;
QY	1	VRVVPVLQQLPQNFQSQQFQREAVPLVQQQCFEPQDQDQFPPOQFYPQDPFFSQQPYLQLQP	60						
DB	21	VRFPVPLQQLPQNFQSQQLPQEQVPLVQQQQLFGQQQFPFPQDPYPQDPFFSQQLPYLQLQP	80						
QY	61	FPQFWLFPQDQSFPPQDPFYPQDPQYSQDQPISSQDQAQQQQQQQQQQQQQQQLLQQLLQ	120						
DB	81	FPQQLPYSQDPFRPQDPYPQDPQYSQDQPISSQDQAQQQQQQQQQQQQQQQLLQQLLQ	140						
QY	121	QQLPQCMQDVVLQENIHARSGQVLCQSTYQLLQELCCQHLMOIPESQCCQAIHNVYHAI	180						
DB	141	QQLPQCMQDVVLQENIHARSGQVLCQSTYQLLQELCCQHLMOIPESQCCQAIHNVYHAI	200						
QY	181	LHQQKQKQKQSSQVSFQDPLQDQDPLQDQGSFRPQDQNPQAQGSVQDQQLQFEEIRNAL	240						
DB	201	LHQQKQKQKQSSQVSFQDPLQDQDPLQDQGSFRPQDQNPQAQGSVQDQQLQFEEIRNAL	260						
QY	241	QTLPAMCNVYIAPYCTIAPFGIFGTN	266						
DB	261	QTLPAMCNVYIAPYCTIAPFGIFGTN	286						

ID	Q9ZP09	PRELIMINARY;	PRT;	288 AA.
AC	Q9ZP09;			
DT	01-MAY-1999 (TREMELrel. 10, Created)			
DT	01-MAY-1999 (TREMELrel. 10, last sequence update)			
DT	01-MAR-2004 (TREMELrel. 26, last annotation update)			
DE	Alpha-gliadin precursor (Fragment).			
GN	Name=alpha-gliadin;			
OS	Triticum aestivum subsp. spelta.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;			
OC	Triticeae; Triticum.			
OX	NCBI_TaxID=58933;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kaarela D.D., D'Ovidio R.;			
RT	"Isolated amino acid sequence of an alpha-gliadin gene from Spelt wheat			
RT	(Spelte) includes sequences active in celiac disease.";			
RL	Cereal Chem. 76:548-551(1999) .			
DR	EMBL; AJ130948; CAA10257.1; .			
DR	PIR; S13333; S13333.			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR003176; Gliadin.			
DR	InterPro; IPR001954; Gliu_glutenn.			
DR	Pfam; PF002344; Tryp_alpha_amy1.			
DR	PRINTS; PR00208; GLIADGLUTEN.			
DR	PRINTS; PR00209; GLIADIN.			
DR	SMART; SM00499; AAI; 1.			
KM	Signal.			
FT	SIGNAL	1	20	Potential.
FT	CHAIN	21	>288	alpha-gliadin.
FT	NON TER	288	288	
SEQ	SEQUENCE	288 AA;	33203 MW;	DA058F3FAFA6BC6C CRC64;
Query Match		94.0%;	Score 1351;	DB 2; Length 288;
Best Local Similarity		95.1%;	Pred. No. 2.6e-75;	
Matches	255;	Conservative	0;	Mismatches 11; Indels 2; Gaps 1;
QY	1	VAVPPOLOQPONPSQQQPOQEVPLVQQQQFPQOQYPPQOPFPQPSQQPYLQLQP	60	
DB	21	VRVPPQLOQPQNPSSQQQPEQVPLVQQQQLGQQQFPQOQYPPQOPFPQPSQQPYLQLQP	80	
QY	61	PPQEWLPFPQPSFPQOQYPPQOQYSSQPPIS--QQQAQQQQQQQQQQQQQQQQQQ	118	
DB	81	PPQPLPVSQPPFPQOQYPPQOQYSSQPPISQQQQQQQQQQQQQQQQQQQQQQQQ	140	
QY	119	LQQQLPCMDVVLQCHNINHARSQLVQSGTYQLLELCCQHLMQIPBQSQCAINNVHA	178	
DB	141	LQQQLPCMDVVLQCHNINHARSQLVQSGTYQLLELCCQHLMQIPBQSQCAIHVVHA	200	
QY	179	ILHQQQKQKQPPSSQVSPQPLQQYPLQGSSFRPSQQNPQAQSVQPPQLPQFEIRKL	238	
DB	201	ILHQQQKQKQPPSSQVSPQPLQQYPLQGSSFRPSQQNPQAQSVQPPQLPQFEIRNL	260	
QY	239	ALQTLPAMCNVYIAPYCTIAPGIRGTN	266	
DB	261	ALQTLPAMCNVYIAPYCTITPFGIRGTN	288	
RESULT 5				
ID	Q9M4M5	PRELIMINARY;	PRT;	274 AA.
AC	Q9M4M5;			
DT	01-OCT-2000 (TREMELrel. 15, Created)			
DT	01-OCT-2000 (TREMELrel. 15, last sequence update)			
DT	01-MAR-2004 (TREMELrel. 26, last annotation update)			
DE	Alpha-gliadin.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;			
OC	Triticeae; Triticum.			

[illegible]

RESULT 9	
ID	PRELIMINARY; PRT; 276 AA.
CM04M0	
CM4M0	
AC	CM4M0;
DT	01-OCT-2000 (TREMblrel. 15, Created)
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT	01-MAR-2004 (TREMblrel. 26, Last annotation update)
DT	Alpha-gladin.

OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OK NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mjcelner; TISSUE=Endosperm;
 RA Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,
 RA Sollid L.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ133608; CAB76960.1; -
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; G1ia-glutenin.
 DR Pfam; PF00234; TYP_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR SMART; SM00499; AAI; 1.
 DR CHAIN 1 276
 FT SEQUENCE 276 AA; 32211 MW; 6A2E9723D42B100A CRC64;
 SQ
 Query Match 93.1%; Score 1338.5; DB 2; Length 276;
 Best Local Similarity 92.7%; Pred. No. 1.5e-74;
 Matches 253; Conservative 2; Mismatches 11; Indels 7; Gaps 1;
 QY 1 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 60
 DB 2 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 61
 QY 61 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 113
 DB 62 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 121
 QY 114 ILQQLIPCMQVPLQGHNIHARSQVLAQSTYQLLQELCCQHLWQIPBOSQCAIHNV 173
 DB 122 ILQQLIPCMQVPLQGHNIHARSQVLAQSTYQLLQELCCQHLWQIPBOSQCAIH 181
 QY 174 NVVHAILHQKQKQSSQVSPFQPLQYPLGSGSFRSPQNPQAGSVQPLQPF 233
 DB 182 NVVHAILHQKQKQSSQVSPFQPLQYPLGSGSFRSPQNPQAGSVQPLQPF 241
 QY 234 EIRNLALQTLPMQCNVIAPYCTIAPFGIFGTN 266
 DB 242 EIRNLALQTLPMQCNVIAPYCTIAPFGIFGTN 274
 RESULT 10
 Q41531 PRELIMINARY; PRT; 289 AA.
 ID Q41531
 AC Q41531
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-gliadin storage protein.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OK NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cheyenne;
 RA Anderson O.D.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U51306; AAA96524.1; -
 DR PIR; S1333; S1333.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; G1ia-glutenin.
 DR Pfam; PF00234; TYP_alpha_amy1; 1.

DR PRINTS; PR00208; GLIADGLUTEN.
 DR PRINTS; PR00209; GLIADIN.
 DR SMART; SM00499; AAI; 1.
 SQ SEQUENCE 289 AA; 33349 MW; 5F577C9CD63874FA CRC64;
 Query Match 92.9%; Score 1334.5; DB 2; Length 289;
 Best Local Similarity 94.1%; Pred. No. 2.7e-74;
 Matches 253; Conservative 1; Mismatches 12; Indels 3; Gaps 1;
 QY 1 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 60
 DB 21 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 80
 QY 61 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 117
 DB 81 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 140
 QY 118 ILQQLIPCMQVPLQGHNIHARSQVLAQSTYQLLQELCCQHLWQIPBOSQCAIHNV 177
 DB 141 ILQQLIPCMQVPLQGHNIHARSQVLAQSTYQLLQELCCQHLWQIPBOSQCAIHNV 200
 QY 178 AIIHQKQKQKQSSQVSPFQPLQYPLGSGSFRSPQNPQAGSVQPLQPF 237
 DB 201 AIIHQKQKQKQSSQVSPFQPLQYPLGSGSFRSPQNPQAGSVQPLQPF 260
 QY 238 LAIQLTLPAMQCNVIAPYCTIAPFGIFGTN 266
 DB 261 LAIQLTLPAMQCNVIAPYCTIAPFGIFGTN 289
 RESULT 11
 Q9M4L9 PRELIMINARY; PRT; 270 AA.
 ID Q9M4L9
 AC Q9M4L9
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-gliadin.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OK NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mjcelner; TISSUE=Endosperm;
 RA Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,
 RA Sollid L.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ133609; CAB76961.1; -
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; G1ia-glutenin.
 DR Pfam; PF00234; TYP_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR SMART; SM00499; AAI; 1.
 FT CHAIN 1 270
 SQ SEQUENCE 270 AA; 31491 MW; 1DB4B6528EADF5 CRC64;
 Query Match 92.4%; Score 1327.5; DB 2; Length 270;
 Best Local Similarity 94.0%; Pred. No. 6.8e-74;
 Matches 251; Conservative 3; Mismatches 12; Indels 1; Gaps 1;
 QY 1 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 60
 DB 2 VRVVPQLQPNPNSQQQPGQVPLVQQQGFPGQQQGFPPQGPYPQPFPSPQPYLQLP 61
 QY 61 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 119
 DB 62 FPGQMLPYPQSGFPFPQGPYPQPFPSPQPYLQLP 121

DR EMBL; U51104; AAA96523.1; -.
DR PIR; S13333; S13333.
DR PIR; T06282; T06282.
DR GO; GO:0045375; F:multient reservoir activity; IEA.
DR InterPro; IPR003512; AA.
DR InterPro; IPR001576; GliaIdn.
DR InterPro; IPR001574; GliaGlutenin.
DR Pfam; PF00234; TYP_alpha_amy; 1.
DR PRINTS; PR00208; GLIADGUTEN.
DR PRINTS; PR00209; GLIADIN.
DR SMART; SM00499; AAT; 1.
SQ SEQUENCE 268 AA; 3328 MW; D7F6B9133283CA2 CRC64;

Query Match	109.1%	DB 2;	length 288;
Best Local Similarity	92.5%	Pred. No. 9.8e-73;	
Matches 248, Conservative	3;	Mismatches 15;	Indels 2; Gaps 1;

Qy	1	VKXVZPOLGQNSQQQPOEQUVILVQKQOPQOQQQOPFPQOPVPOQPFPSQOPYLQOP	60
Db	21	VKXVZPOLGQNSQQQPOEQUVILVQKQOPFLQKQOPFPQOPVPOQPFPSQOPYLQOP	80
Qy	61	FPQWMLTYPQSGFPFPQOPYPOFPQPSQOPQPSQQAQQQQQQQQQQQ--QQ11QQ1	118
Db	81	FSQOLPYSPQPFRRPQOPQOPYQSPQOPQPSQQAQQQQQQQQQQQQQQQQ11QQ1	140
Qy	119	LQOQLIPCMNVVLQOHNIHARSQVLAQSTYLLLELCCOHLMQIPESQCCQALHNHVHA	178
Db	141	LQOQLIPCMNVVLQOHNIHARSQVLAQSTYLLLELCCOHLMQIPETLQCOALHNHVHA	200
Qy	179	IIILHQKQKQOQSSQVSPQQLPQQYPLGQSGFRPSQNPQAQSGVQPPQQLPQEEIRNL	238
Db	201	IIILHQKQKQOQSSQVSPQQLPQQYPLGQSGFRPSQNPQAQSGVQPPQQLPQEEIRNL	260
Qy	239	ALQTLPMQNVYLAQYCTIAPFQIRFTN	266
Db	261	ALQTLPMQNVYLAQYCTIAPFQIRFTN	288

```

RESULT 15
Q41528
ID Q41528 PRELIMINARY; PRT; 287 AA.
AC Q41528;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Alpha-glutinin.
OS Triticum aestivum (Wheat) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; Liliopsida; Poales; Poaceae
OC Gramineae; Triticaceae; Triticum.
OX NCBI_TaxID=4565;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cheyenne;
RA Anderson O.D.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases
DR EMBL; U50984; AAA86276.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IFA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001376; Glia.
DR InterPro; IPR001954; Glia glutenin.
DR Pfam; PF00234; Tryp alpha amyl; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR SMART; SM00209; GLIADIN.
DR SMART; SM00399; AAI; 1..
SQ SEQUENCE 287 AA; 33193 MW; 05F82296749C9E87 CRC64;

```

Query Match	90.2%	Score 1296.5	DB 2	Length 287
Best Local Similarity	92.9%	Pred. No. 5.7e-72		
Matches 249; Conservative	2	Mismatches 14	Indels 3	Gaps 3

QY 1 V R V E P P Q L P Q N F S Q Q Q P Q D E V P L V Q Q Q Q P F G Q Q Q Q P P P Q Q P Y P Q Q P P E S Q Q P Y L Q L Q P 600

Db 2 VRFPPVQLQPNPSSQQLPQECVGLVQKQQLFLGQQQQPFPPQCPYFQPC-FESQLPYLQLP 76

Qy 61 FFPPLPYRQGSFPPQCPYRQPCQYSPQPCPS-00A0Q0Q0Q0Q0Q0Q0Q0Q0QLLOQL 119

Db 80 FPPQLPYQPQPFRRQCPYRQPCQYSPQPCPSQ000000000000000000QLLOQL 139

Qy 120 QQQLLPCMDVVLQOHNIHAARSQVLQOSTYLLQELCCQHLMOIPEQSCCAINNVHAI 179

Db 140 QQQLLPCMDVVLQOHNHKAGRSQVLQOSTYLLRRLCCQHLMOIPEQSCCAINNVHAI 199

Qy 180 ILHQQR-QQQPSSQYSPQLQYPRFGQSSFPSSQNNPQAQSVQPCQLPQHEETRL 239

Db 200 ILHQQRQ000Q0PSSQVSFQQLQ0YPRFGQSSFPSSQNNPQTQSVQPCQLPQHEETRL 259

Qy 239 ALQTLPMGNVYIAPCTIAPFGIRGTN 266

Db 260 ALQTLPMGNVYIIPYCTIAPFGIRGTN 287

Search completed: December 14, 2004, 17:25:23
Job time : 80 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:00:05 ; Search time 76.6667 Seconds
(without alignments)
1244.635 Million cell updates/sec

Title: US-10-089-700-3-Y65

Perfect score: 1433

Sequence: 1 VRVPPQLQPNPSCQCPQF.....CNVTIAPYCTIAPRGIGTN 266

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq23Sep04:*
1: geneeqp1980s:*
2: geneeqp1990s:*
3: geneeqp2000s:*
4: geneeqp2001s:*
5: geneeqp2002s:*
6: geneeqp2003as:*
7: geneeqp2003bs:*
8: geneeqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1425	99.4	266 4	AAU01799
2	1425	99.4	266 8	ADH14513
3	1265.5	88.3	290 7	AAE38574
4	1265.5	88.3	290 8	ADP19626
5	445	32.4	369 2	AAW62647
6	450.5	31.4	287 8	ADOT1669
7	438.5	30.6	307 7	ADH89358
8	437.5	30.5	307 8	ADG44134
9	437.5	30.5	298 8	ADOT1661
10	285.5	19.9	1798 4	ABR71695
11	269	18.8	2285 4	ABR63057
12	266	18.6	1162 3	AAV96255
13	266	18.6	1162 3	AAV85800
14	266	18.6	1162 4	AAE2331
15	266	18.6	1162 5	ABR05621
16	266	18.6	1162 8	ADJ56096
17	264.5	18.5	757 8	ADJ56096
18	264.5	18.5	1069 6	ABO07138
19	264.5	18.5	1069 6	ADJ37233
20	255.5	17.8	260 8	ADOT4673
21	253	17.7	186 8	ADH89336
22	253	17.7	186 8	ADG44132
23	253	17.7	905 5	ABG93053
24	253	17.7	905 6	ABR53130
25	253	17.7	905 7	ADK62564

ALIGNMENTS

26	247.5	17.3	900	4	ABR62018	AbB62018 Drosophil
27	244.5	17.1	1069	4	ABR61305	AbB61305 Drosophil
28	244	17.0	1013	4	ABR71093	AbB71093 Drosophil
29	239	16.7	358	7	ADH65556	ADH65556 Human pro
30	236.5	16.5	1142	7	ADCO7968	ADCO7968 Rice prot
31	231.5	16.2	149	4	ADH89335	ADH89335 H. vulgar
32	231	16.1	160	7	ADH89335	ADH89335 H. vulgar
33	231	16.1	160	8	ADG44131	ADG44131 H. vulgar
34	231	16.1	160	8	ADG44131	ADG44131 H. vulgar
35	230	16.1	160	8	ADG44131	ADG44131 H. vulgar
36	230	16.1	160	8	ADG44131	ADG44131 H. vulgar
37	230	16.1	160	8	ADG44131	ADG44131 H. vulgar
38	228	15.9	158	3	AAV54568	AAV54568 A synthe
39	228	15.9	2280	4	ABR61650	ABR61650 Drosophil
40	226	15.8	153	3	AAV69495	AAV69495 Amino aci
41	226	15.8	738	5	ABG93140	ABG93140 S. cerevi
42	224	15.6	1428	4	ABR70377	ABR70377 Drosophil
43	223.5	15.6	785	8	ADP98983	ADP98983 C. albica
44	223	15.6	368	4	ABR63167	ABR63167 Drosophil
45	222.5	15.5	467	8	ADJ76333	ADJ76333 Marker ge

RESULT 1
AAU01799 standard; protein; 266 AA.

AAU01799
ID AAU01799 standard; protein; 266 AA.
AC AAU01799;
AD 07-SEP-2001 (first entry)
AE Wheat A-gliadin.
AF Wheat A-gliadin.
AG Wheat A-gliadin; epitope; coeliac disease; gluten intolerance;
AH T-cell binding; antagonist; transglutaminase; transgenic plant.
AI Triticum aestivum.
AJ WO200125793-A2.
AK 12-APR-2001.
AL 02-OCT-2000; 2000WO-GB003760.
AM 01-OCT-1999; 99GB-00023306.
AN (ISIS-) ISIS INNOVATION LTD.
AO Anderson RP, Hill AVS, Jewell DP;
AP WPI; 2001-300179/31.
AQ
AR
AS
AT
AU
AV
AW
AX
AY
AZ
BA
BB
BC
BD
BE
BF
BG
BH
BI
BJ
BK
BL
BM
BN
BO
BP
BQ
BR
BS
BT
BU
BV
BW
BX
BY
BZ
CA
CB
CC
CD
CE
CF
CG
CH
CI
CJ
CK
CL
CM
CN
CO
CP
CQ
CR
CS
CT
CU
CV
CW
CX
CY
CZ
DA
DB
DC
DD
DE
DF
DG
DH
DI
DJ
DK
DL
DM
DN
DO
DP
DQ
DR
DS
DT
DU
DV
DW
DX
DY
DZ
EA
EB
EC
ED
EE
EF
EG
EH
EI
EJ
EK
EL
EM
EN
EO
EP
EQ
ER
ES
ET
EU
EV
EW
EX
EY
EZ
FA
FB
FC
FD
FE
FF
FG
FH
FI
FJ
FK
FL
FM
FN
FO
FP
FQ
FR
FS
FT
FU
FV
FW
FX
FY
FZ
GA
GB
GC
GD
GE
GF
GG
GH
GI
GJ
GK
GL
GM
GN
GO
GP
GQ
GR
GS
GT
GU
GV
GW
GX
GY
GZ
HA
HB
HC
HD
HE
HF
HG
HH
HI
HJ
HK
HL
HM
HN
HO
HP
HQ
HR
HS
HT
HU
HV
HW
HX
HY
HZ
IA
IB
IC
ID
IE
IF
IG
IH
II
IJ
IK
IL
IM
IN
IO
IP
IQ
IR
IS
IT
IU
IV
IW
IX
IY
IZ
JA
JB
JC
JD
JE
JF
JG
JH
JI
JJ
JK
JL
JM
JN
JO
JP
JQ
JR
JS
JT
JU
JV
JW
JX
JY
JZ
KA
KB
KC
KD
KE
KF
KG
KH
KI
KJ
KK
KL
KM
KN
KO
KP
KQ
KR
KS
KT
KU
KV
KW
KX
KY
KZ
LA
LB
LC
LD
LE
LF
LG
LH
LI
LJ
LK
LM
LN
LO
LP
LQ
LR
LS
LT
LU
LV
LW
LX
LY
LZ
MA
MB
MC
MD
ME
MF
MG
MH
MI
MJ
MK
ML
MN
MO
MP
MQ
MR
MS
MT
MU
MV
MW
MX
MY
MZ
NA
NB
NC
ND
NE
NF
NG
NH
NI
NJ
NK
NL
NM
NO
NP
NQ
NR
NS
NT
NU
NV
NW
NX
NY
NZ
OA
OB
OC
OD
OE
OF
OG
OH
OI
OJ
OK
OL
OM
ON
OO
OP
OQ
OR
OS
OT
OU
OV
OW
OX
OY
OZ
PA
PB
PC
PD
PE
PF
PG
PH
PI
PJ
PK
PL
PM
PN
PO
PP
PQ
PR
PS
PT
PU
PV
PW
PX
PY
PZ
QA
QB
QC
QD
QE
QF
QG
QH
QI
QJ
QK
QL
QM
QN
QO
QP
QQ
QR
QS
QT
QU
QV
QW
QX
QY
QZ
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ
SA
SB
SC
SD
SE
SF
SG
SH
SI
SJ
SK
SL
SM
SN
SO
SP
SQ
SR
SS
ST
SU
SV
SW
SX
SY
SZ
TA
TB
TC
TD
TE
TF
TG
TH
TI
TJ
TK
TL
TM
TN
TO
TP
TQ
TR
TS
TT
TU
TV
TW
TX
TY
TZ
UA
UB
UC
UD
UE
UF
UG
UH
UI
UJ
UK
UL
UM
UN
UO
UP
UQ
UR
US
UT
UU
UV
UW
UX
UY
UZ
VA
VB
VC
VD
VE
VF
VG
VH
VI
VJ
VK
VL
VM
VN
VO
VP
VQ
VR
VS
VT
VU
VV
VW
VX
VY
VZ
WA
WB
WC
WD
WE
WF
WG
WH
WI
WJ
WK
WL
WM
WN
WO
WP
WQ
WR
WS
WT
WU
WV
WW
WX
WY
WZ
XA
XB
XC
XD
XE
XF
XG
XH
XI
XJ
XK
XL
XM
XN
XO
XP
XQ
XR
XS
XT
XU
XV
XW
XX
XY
XZ
YA
YB
YC
YD
YE
YF
YG
YH
YI
YJ
YK
YL
YM
YN
YO
YP
YQ
YR
YS
YT
YU
YV
YW
YX
YY
YZ
ZA
ZB
ZC
ZD
ZE
ZF
ZG
ZH
ZI
ZJ
ZK
ZL
ZM
ZN
ZO
ZP
ZQ
ZR
ZS
ZT
ZU
ZV
ZW
ZX
ZY
ZZ

Claim 1; Page 52; 107p; English.

The sequence represents wheat A-gliadin. A-gliadin derived peptides of the invention are used to test mammalian (preferably human) susceptibility to coeliac disease (gluten intolerance). The peptides are contacted with a blood sample and T cell recognition measured, a positive T-cell recognition indicating a susceptibility to coeliac disease. The peptides are useful for inducing tolerance in an individual and for producing an antibody or preventing coeliac disease in an individual and for producing an antibody specific to them or a wild-type sequence. A mutant gliadin protein (or its fragment of 15 amino acids in length) whose wild-type sequence can be modified by transglutaminase to a sequence that comprises the epitope, but which has been modified in such a way that it does not contain sequence which can

GC be modified by transglutaminase to a sequence that comprise the epitope
 GC is useful for decreasing the ability of gliadin protein to cause Coeliac
 GC disease. Nucleic acids encoding proteins antagonistic to the T-cell
 GC binding of the epitopes are useful for obtaining a transgenic plant cell
 GC or seed and for the production of a protein. The resultant crop plant is
 GC useful for obtaining a product of a wheat plant, especially grain, which
 GC is optionally processed into flour or another grain product. Food
 GC comprising the antagonistic protein is useful instead of a wild-type
 CC gliadin

XX Sequence 266 AA;

XX Query Match

Best Local Similarity 99.4%; Score 1425; DB 4; Length 266;

Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VRVVPVQLQPNPSSQVQPEQVPLVQQQPFPGQQQPFPPQPPYPQPPFSQCPYTLQLP 60
 1 VRVVPVQLQPNPSSQVQPEQVPLVQQQPFPGQQQPFPPQPPYPQPPFSQCPYTLQLP 60
 61 FPGVPLPFPQPSFPQPPQPPYQPPQYSPQPPISQQAQQQQQQQQQQQQQQQQQQQLLQQLLQ 120
 61 FPGVPLPFPQPSFPQPPQPPYQPPQYSPQPPISQQAQQQQQQQQQQQQQQQQQQQLLQQLLQ 120
 121 QQLIPCMQDVVLQCHNIAHARSQVLAQSTYQLCLCELCCHLMQIPBQSCCAIHNVVAII 180
 121 QQLIPCMQDVVLQCHNIAHARSQVLAQSTYQLCLCELCCHLMQIPBQSCCAIHNVVAII 180
 181 LHQQKXQQQPSQVSPQPPQYPLQGGSFRRPSCQNPQAGSVQPPQLPQFEIRNLAL 240
 181 LHQQKXQQQPSQVSPQPPQYPLQGGSFRRPSCQNPQAGSVQPPQLPQFEIRNLAL 240
 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266
 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266

RESULT 2

ADH14513

ID ADH14513 standard; protein; 266 AA.

XX ADH14513;

XX 11-MAR-2004 (first entry)

XX A-gliadin protein sequence SEQ ID NO:3.

XX coeliac disease; gliadin; gliadin T cell epitope; gastrointestinal;

XX vaccine.

XX Synthetic.

XX MO2003104273-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-GB002450.

XX 05-JUN-2002; 2002GB-00012885.

XX (ISIS-) ISIS INNOVATION LTD.

XX Anderson RP, Hill AVS, Jewell DP;

XX WPT; 2004-043640/04.

XX Preventing or treating coeliac disease comprises administering agent

XX which are wheat gliadin T cell epitope capable of being recognized by T

XX cell receptor.

XX Example 1; SEQ ID NO 3; 177pp; English.

XX The present invention describes a method (M1) for preventing or treating

CC coeliac disease. M1 comprises administering an agent (A) comprising a
 CC gliadin T cell epitope which is capable of being recognised by a T cell
 CC receptor, to an individual. Gliadin is a component of gluten. (A) has
 CC gastrointestinal activity, and can be used in vaccines. The agent (A) can
 CC be used in the preparation of a medicament for treating or preventing
 CC coeliac disease. (A) can also be used in the preparation of a diagnostic
 CC means for use in diagnosing coeliac disease, or susceptibility to coeliac
 CC disease, in an individual, which involves determining whether T cells of
 CC the individual recognise the agent, recognition by the T cells indicating
 CC that the individual has, or is susceptible to, coeliac disease. The
 CC present sequence represents a protein which is used in the
 CC exemplification of the present invention.

XX Sequence 266 AA;

XX Query Match

Best Local Similarity 99.4%; Score 1425; DB 8; Length 266;

Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VRVVPVQLQPNPSSQVQPEQVPLVQQQPFPGQQQPFPPQPPYPQPPFSQCPYTLQLP 60
 1 VRVVPVQLQPNPSSQVQPEQVPLVQQQPFPGQQQPFPPQPPYPQPPFSQCPYTLQLP 60
 61 FPGVPLPFPQPSFPQPPQPPYQPPQYSPQPPISQQAQQQQQQQQQQQQQQQQQQQLLQQLLQ 120
 61 FPGVPLPFPQPSFPQPPQPPYQPPQYSPQPPISQQAQQQQQQQQQQQQQQQQQQQLLQQLLQ 120
 121 QQLIPCMQDVVLQCHNIAHARSQVLAQSTYQLCLCELCCHLMQIPBQSCCAIHNVVAII 180
 121 QQLIPCMQDVVLQCHNIAHARSQVLAQSTYQLCLCELCCHLMQIPBQSCCAIHNVVAII 180
 181 LHQQKXQQQPSQVSPQPPQYPLQGGSFRRPSCQNPQAGSVQPPQLPQFEIRNLAL 240
 181 LHQQKXQQQPSQVSPQPPQYPLQGGSFRRPSCQNPQAGSVQPPQLPQFEIRNLAL 240
 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266
 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266

RESULT 3

AAE38574

ID AAE38574 standard; protein; 290 AA.

XX AAE38574;

XX 04-DEC-2003 (first entry)

XX Wheat alpha-2 gliadin protein.

XX Wheat; therapy; celiac sprue; dermatitis herpetiformis; gluten toxicity;

XX glutenase; foodstuff; antiinflammatory; dermatological; alpha-2 gliadin.

XX Triticum aestivum.

XX WO2003068170-A2.

XX 21-AUG-2003.

XX 14-FEB-2003; 2003WO-US004743.

XX 14-FEB-2002; 2002US-0357238P.

XX 14-MAY-2002; 2002US-0380761P.

XX 28-JUN-2002; 2002US-0392782P.

XX 31-OCT-2002; 2002US-0432933P.

XX 20-NOV-2002; 2002US-0428033P.

XX 20-DEC-2002; 2002US-0435881P.

XX (STRD) UNIV LEIAND STANFORD JUNIOR.

XX Hausch F, Gray G, Shan L, Khosla C;

XX WPT; 2003-697466/66.

XX The present sequence represents the mature glutenin protein. The DNA
CC sequence encoding this protein is isolated from the genomic DNA of
CC Triticum durum L. The gene codes for a low-molecular-weight glutenin
CC protein and can be used to produce transgenic durum wheat plants with
CC "better quality characteristics" (no details given). (Updated on 25-MAR-
CC 2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS
CC field)
SQ Sequence 369 AA;
Query Match 32.4%; Score 465; DB 2; Length 369;
Best Local Similarity 40.2%; Pred. No. 5.6e-32;
Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;
6 POLQPNSQQQP-----QEQVPLVQGGQPPGQQQPP-----PQPPYPPQPPPP-----50
38 PQQPPCQQQQQQPPPSQQQQPPPSQQQQPPPSQQQQPPPSQQQQPPPSQQQQPP-----97
51 ---SQQPYL-----QLOPPQPYLPYPPQPPPPQ-----PYPPQPPQYQ 83
98 FQQQQQPPVLPQPPPSQQQLPPPSQQQLPPPSQQQLPPPSQQQLPPPSQQQLPPPSQ 157
90 PQQPSQQQQAQQQQQQQ-----QQQQQQQILQQ-----ILQQQLIPQMDVVLQ 132
158 QQQVVLVQQPPPSQQQQPP 216
133 QH-----NIAHARSQVLAQSTYQLQELCCQHLWQIPESQCCQAIHNVVHAILHQQK 186
217 QQQSPWMPQSLARSCMLQQSSCHWQCCQQLPQIPQPSRYEALRAIVYSIIL--QEQ 274
187 QQQQPSQVSRQQLQGYPRGQSGFRPSQGNPQAQS-----VQPPQLPPFPE 234
275 QQQVGSQTQQQQQPPQ---LQQVSPQQQSQQQQLGQQPQQQLAHGTFLQPPQIMLEY 331
235 IRNLALQTLPAMCNVYIAPY--CTIAPFGIFGT 265
332 MTSIALRTPLTMCNWNVPLVTRTTRVPGV--GT 363
RESULT 6
AD071669
ID AD071669 standard; protein; 297 AA.
XX
AG AD071669;
XX
DE 12-AUG-2004 (first entry)
XX
DE Amino acid sequence of a modified glutenin LMW subunit.
XX
KW low molecular weight subunit; LMW subunit; glutenin;
KW wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;
KW gluten intolerance.
XX
OS Triticum sp.
OS Synthetic.
XX
EV BP1424342-A1.
XX
DE 02-JUN-2004.
XX
DE 27-NOV-2002; 2002EP-00026461.
XX
DE 27-NOV-2002; 2002EP-00026461.
XX
DE (BAKE-) BAKEMARK DEUT GMBH.
DE (MONS) MONSANTO AGRAR DEUT GMBH.
DE (UNIF-) UNIFERN GMBH & CO KG.
DE (PURA-) PURATOS NV.
XX
DE Hinzmann E, Wieser H, Stahl U;
XX
DE MPI; 2004-402870/38.

DR N-PSDB; AD071668.
XX
XX Novel nucleic acid comprising sequence encoding modified glutenin
PT polypeptide, useful for preparing modified glutenin polypeptide as
PT gliadin substitute in foodstuffs such as dough, pastries and wafers.
XX
XX Claim 16; Fig 11; 43pp; English.
XX
XX The present sequence represents a modified low molecular weight (LMW)
CC subunit of glutenin. The wild type subunit is designated clone LMW6, and
CC is isolated from wheat cultivar Cheyenne. The LMW6 polypeptide does not
CC contain the allergenic epitope QQQPP, and shows some minor differences to
CC published sequences. It therefore represents a new allele for LMW subunit
CC genes. The LMW6 polypeptide was modified to produce modified glutenin
CC polypeptides of the invention. In these modified polypeptides one or more
CC cysteine residues responsible for intermolecular cross linking through
CC disulfide bridges are deleted or substituted. The modified glutenin
CC polypeptide is useful as a gliadin substitute. It is also useful in the
CC preparation of foodstuffs, such as flour or for the preparation of
CC pharmaceutical products, such as tablets, where the foodstuffs contain a
CC considerably reduced amount of gliadin proteins or no gliadin proteins.
CC Pharmaceutical compositions comprising the modified polypeptide of the
CC invention are useful for treating patients suffering from coeliac disease
CC or persons who are intolerant to gluten.
XX
SQ Sequence 297 AA;
Query Match 31.4%; Score 450.5; DB 8; Length 297;
Best Local Similarity 41.2%; Pred. No. 7.9e-31;
Matches 120; Conservative 38; Mismatches 82; Indels 51; Gaps 13;
5 VPQLQPNSQQQPPQPPVPLVQGGQPPGQQQPPPPPPPPPPPPPPPPPPPPPPPPPPPP 63
18 IQMETSIPGSEFWQQQPLQKXTFP---QPPSSQ---QQPPFPQPPFLQQPPPSFSQ 71
64 PVLPPQPSPP 116
72 -----QLPFGQQQPPVLPQPPPSQQQQLPQPPPSQQQQLQQLQQLQQLQQLQQLQ 125
117 QI-----LQQQLIPQMDVVLQGNINIAHARSQVLAQSTYQLQELCCQHLWQIPESQCC 170
126 QLNPKVFLQQQ---CSPVAMPQH---LARSQMWQSSCNVMQCCQQLPRLPESQRYE 173
171 AIHNVVHAILIHCQOK-----QQQPPSSQV--SFQQPLQ---QYPLQGQSF-----RP 213
180 AIPRAITFILLQEQQGGVGPQQQQPPQSGVGVYPQPPQSGQQLGQSGFPQQPQQQLGQPF 239
214 SQQNPQAQGSVQPPQLPPFESIRMLAQTLPAMCNVYIAPYCTI--APPFG 262
240 QQQQVQKGFLLQPPHIALLEVMTSIALRTPLTMCQSVNVPYISITSAPLGV 290
RESULT 7
ADH89338
ID ADH89338 standard; protein; 307 AA.
XX
AC ADH89338;
XX
DT 06-MAY-2004 (first entry)
XX
DE T. aestivum LMW glutenin-1D1 protein.
XX
XX double stranded RNA; storage protein; 2S-albumen; 7S-globulin;
KW 11S/12S-globulin; zein-prolamine; homogenistate metabolic pathway;
KW pharmaceutical; plant; abiotic stress; fatty acid composition;
KW lipid composition; oil composition; carbohydrate composition; colour;
KW pigmentation; pathogen resistance; fruit ripening delay; aging;
KW male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;
KW caffeine; theophylline; threonine biosynthesis; glutenin.
XX
OS Triticum aestivum.
OS
XX
XX WO2003078629-A1.

Pd		XX	25-SEP-2003.		
Xx		PF	17-MAR-2003; 2003WO-EPO02735.		
Pf		PR	20-MAR-2002; 2002DE-01012892.		
Xx		PA	(BADI) BASF PLANT SCI GEMBH.		
Pa		PI	Kock M., Bauer J;		
Dt		DR	NFI; 2003-603869/75.		
Rn		DR	N-PSDB; ADH89337.		
Pt		PT	Reducing expression of at least two target genes, useful e.g. for producing transgenic plants, using partly double-stranded interfering RNA..		
Fs		FS	Disclosure, SEQ ID NO 113; 228pp; German.		
Xx		XX	This invention describes a novel method for reducing the expression of at least two different endogenous target genes in a eukaryotic cell or organism by introducing an RNA molecule that is at least partly double stranded. The transcribed RNAs form at least two target genes have homology below 90% and the RNA molecule is formed as a single, self-complementary molecule. At least one of the double-stranded structures formed from individual sense sequences has an even number of repeats of CC 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At least two target genes are selected from different classes of storage protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-prolamine and at least one of the sense sequences is identical to storage protein sequences or genes in the homogentisate metabolic pathway or enzyme types, e.g. acetyl transferases, cholinesterase, (de)branching enzymes or cellulases. The RNA of the invention, also related cassettes, expression systems, vectors and transgenic organisms are used for preparation of pharmaceuticals, in biotechnological processes and plant biotechnology specifically in plants to improve protection against abiotic stress, to modify composition and/or content of fatty acids, lipids and oils, to modify carbohydrate composition, to alter colour or pigmentation, to reduce content of storage proteins, to increase resistance to pathogens, to inhibit stem break, to delay fruit ripening or aging, to induce male sterility, to reduce content of toxic or unwanted components, to modify lignification and/or lignin content, to modify the fibre component in foods or fibre quality in cotton, to reduce susceptibility to shock, to increase synthesis of Vitamin E, to reduce contents of nicotine, caffeine or theophylline and to increase methionite content, by reducing threonine biosynthesis. The method provides a rapid and efficient way of reducing gene expression, can inhibit more than one target gene, prevents development of multiple phenotypes (since the transcription rate is the same for all RNA sequences, significantly reducing the selection process required to produce an organism with effective suppression of all target genes), avoids problems of epigenic gene silencing, does not require synthesis of individual RNA sequences and the method can be applied to plants with complex (polyploid) genomes. No interference between the individual RNA sequences occur. This sequence represents a protein encoded by a target gene used in the method of the invention.		
Sq		XX	Sequence 307 AA:		
Qy	Query Match		30.6%; Score 438.5; DB 7; Length 307;		
Bb	Best Local Similarity		42.8%; Pred. No. 9e-30;		
Matches	Matches 125; Conservative 36; Mismatched 76; Indels 55; Gaps 16				
Yy	13 P S Q Q Q A D G V I N Y C C Q F P-----G G Q Q Q F P P Q Q R P V Q G Q P P L T L Q A P P P O Y L R 67 ::: 27 P G L R P W Q Q Q L P P Q Q F P Q Q P L F S Q Q Q---C Q L F P Q Q P S F S Q Q P-----P W Q Q Q P P 78				
Oy	68 Y P Q P S P F P P Q P-----Y P Q P Q P O Y S P Q P C P I-----S Q Q A C Q Q Q Q Q Q Q Q Q Q Q Q I- 114 : 79 F S Q Q Q P L P Q Q P P P S Q Q Q A V L H Q-Q P P S Q Q Q Q Q P V L P P Q G P F P F Q Q Q Q Q Q H Q Q-V Q Q Q I P 137				
Zz	115 -L Q Q L L Q Q Q L I P C K D V T L Q Q H----N I A H-----A R S Q V L Q Q S T Y Q L L Q L C C G H M Q I P B Q S 167				

Db 138 VQGSITLQQLNPFCKVFLVQGGCCSVNMPQRLANSQMLQSSCHVMQGGCCQLPQIRQGS 196
 Qy 168 QCCAIHNHVAIIILHQGQKQGGQSSGVSFQQLPQVPLQGGSRPSQGNFQ----- 219
 Db 197 RYEAIRLRIYSIIL--QEGCGQVGSISQGGCGPQQ--LQGVSGPQGGQSGQULGGQPPQ 251
 Qy 220 ---AGSG-VQPGQLPGFEETRNIALQTLPMVCNVIAPY--CTIAPFGIFGT 265
 Db 252 QQLAGTFFLPQHPQIAQLQLEVMFTSILRLPLPMCSNVNPLRTTTSVPGGV-CT 302
 RESULT 8
 ID ADG44134 standard; protein; 307 AA.
 AC ADG44134;
 XX
 DT 26-FEB-2004 (first entry)
 DE T. aestivum glutenin-1D1 protein.
 XX oil content; plant; storage protein; seed-specific promoter; 28-albumin;
 KM 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;
 KM oil production; fat production; free fatty acid production; food;
 KM animal feed; pharmaceutical; fine chemical production; glutenin.
 XX
 OS Triticum aestivum.
 XX
 PN MO200307643-A2.
 PD 25-SEP-2003.
 XX
 PF 17-MAR-2003; 2003WO-EP002733.
 PR 20-MAR-2002; 2002DE-01012893.
 XX
 PA (BADI) BASF PLANT SCI GMBH.
 PI Bauer J;
 DR WPI; 2004-011485/O1.
 DR N-PSDB; ADG44133.
 XX
 PT Increasing total oil content of plants, useful e.g. as foods or animal
 PT feeds, by reducing amount of storage proteins, particularly with double-
 PT stranded interfering RNA.
 PS
 PS Claim 4; SEQ ID NO 174; 253bp; German.
 XX
 XX This invention describes a novel method for increasing the total oil
 XX content of a plant by reducing the amount of at least one storage protein
 XX in the plant (or its tissue, organs, parts or cells) and selecting plants
 XX that have higher total oil content than starting plants. The storage
 XX protein is suppressed by introducing antisense RNA, optionally combined
 XX with a ribozyme, sense RNA that induces co-suppression, DNA-binding
 XX factors directed against storage protein genes, viral sequences that
 XX degrade storage protein RNA, constructs that induce homologous into
 XX recombination of endogenous storage protein genes or mutations into
 XX storage protein genes. Most preferably a plant cell is stably transfected
 XX with a recombinant expression construct, then regenerated to plants that
 XX express the incorporated sequence. The expression constructs particularly
 XX contain a seed-specific promoter and they are introduced into plants by
 XX standard methods, e.g. via Agrobacterium. The preferred storage proteins
 XX of the invention are 2S-albumens, 7S or 11S/12S-globulins or zein-
 XX prolamines. Transgenic organisms produced by the new method are used for
 XX production of oils, fats, free fatty acids or their derivatives, useful
 XX as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence
 XX represents a storage protein used to illustrate the method of the
 XX invention.
 XX
 XX Sequence 307 AA:
 XX

Query	Subject	Score	Length	Start	End	Ident	Score	Length	Start	End	Ident
1	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
2	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
3	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
4	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
5	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
6	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
7	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
8	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
9	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
10	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
11	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
12	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
13	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
14	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
15	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
16	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
17	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
18	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
19	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
20	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
21	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
22	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
23	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
24	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
25	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
26	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
27	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
28	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
29	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
30	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
31	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
32	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
33	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
34	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
35	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
36	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
37	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
38	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100%
39	AB071695	41.0%	298	1	298	100%	41.0%	298	1	298	100

DR N-PSDB; ABL15798.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 41877; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
 CC AB372072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 1798 AA;
 XX
 Query Match 19.9%; Score 285.5; DB 4; Length 1798;
 Best Local Similarity 37.3%; Pred. No. 1.3e-15;
 Matches 100; Conservative 18; Mismatches 93; Indels 57; Gaps 12;
 QY 3 VPVPQLQPNP-SQQQPEQVPLVQ-----QQQPPQQQQPPQQ 41
 Db 264 VEGQATQPFQSQCKFIDPTDPCVAVLSALSNDLSLIMRQQLKQQQMQQQQQ 323
 QY 42 PYPQPPFP-SQQPVLQLPFPQFVLPYQPOSFPQQ-----PYPQPPQVYSQPPQPSQ 96
 Db 324 WAPQPPQQAQPPQQQQQQPPQQ-QHTSPQSPFQQQPTPTLQQQPPNQMAQQ-1QQ 381
 QY 97 QQAQQQQQQQQQQQQQQQLLQQLLQQLLP-CMDVVLQGH-NIAHBSQVLQGSTYQLL 152
 Db 382 QQQQQQQQQQQQQQQQQQQQLLQQLLPQPCQQQQQVITGRHVINTSTAGQQQLIQSHMSL- 440
 QY 153 QELCCGHLMOIPBQSCCAIHNVVAILIHQQQKQQQSSQVSTQPPQLQVPLQGSFPR 212
 Db 441 -----ALQKQQQ-----LHVQQQAQQQPPQQQQQLTQVQLPFAQQQQQL 480
 QY 213 PQQQNPQAQSVQPPQLPQFEIRNLAL 240
 Db 481 PQGHVQQQ---QPPQV-QFTQQQQAL 504

RESULT 11
 ABB63057
 ID ABB63057 standard; protein; 2285 AA.

XX ABB63057;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polyprotein SEQ ID NO 15963.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams N, Li FWD, Myers EW,
 FI

XX WPI; 2001-656860/75.
 DR N-PSDB; ABL07160.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 15963; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
 CC AB372072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 2285 AA;
 XX
 Query Match 18.8%; Score 269; DB 4; Length 2285;
 Best Local Similarity 35.7%; Pred. No. 4.5e-14;
 Matches 101; Conservative 14; Mismatches 104; Indels 64; Gaps 13;
 QY 7 QLPQNP-----SQQPEQVPLVQQQFPQQQQPPQQPPQPPQ-----PFP-SQQP 54
 Db 909 QMQQQQVVAAPVTHNQVMPQQQVN--QQQQQPMQQLPQQVQVQBPVLPFPQPHQDP 966
 QY 55 YLQHPFPQ-----PYLPYQPOSFPQQFPYQPPQVYSQPPQPSQQA--- 99
 Db 967 QQQQPPQLQLQLMHTNVQAPDL--TQQQMAQQQAQVYFQQQQQQQPPQAVNMQQAAYAM 1024
 QY 100 QQQQQQQQQQQQQQLLQQLLQQLI-----PCMDVVLQGHNIAHBSQVLQGSTYQLL 153
 Db 1025 QQAQQQQQLPLQLLQQLLQQLLQQLQQAQVAVSHQQQIMQQQLAQHQLQQLLQQLLQQLQ 1084
 QY 154 ELCCGHLMOIPBQSCCAIHNVVAILIHQQQKQQQPPSSQV-----SQQPQLQ-----QY 204
 Db 1085 QIQQQQLQQLQLQQLQQLQ-QFVQQYACA-MPQQQHQQLVTSQVWAFHQHQPPQLQPVQM 1141
 QY 205 PL-----GQGSFPPSQQPPQAQGSVPQQLPQF 232
 Db 1142 PFTSVAPPQHTYVQGGGVTLSDAQQQQHPPGFAVQQAAPF 1184

RESULT 12
 AAY96255
 ID AAY96255 standard; protein; 1162 AA.

XX AAY96255;
 AC
 XX 12-SEP-2003 (revised)
 DT
 XX 11-SEP-2000 (first entry)
 DT
 XX Kaposi's sarcoma-associated herpesvirus LANA.
 DE
 XX Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
 KW latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
 KW Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
 KW Kaposi's sarcoma; primary effusion lymphoma; PEL;
 KW human immunodeficiency virus; HIV; multicentric Castleman's disease.
 XX
 XX Human herpesvirus 8.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 14..17
 FT /note="nuclear localisation signal, NLS"
 FT Domain 64..70
 FT /note="nuclear localisation signal, NLS"
 FT Region 320..429

Best Local Similarity 38.0%; Pred. No. 3.8e-14;

Query Match	18.6%;	Score 256;	DB 3;	Length 1162;
Best Local Similarity	38.0%;	Pred. No. 3.8e-14;		
Matches	93;	Conservative	19;	Mismatches 101; Indels 32; Gaps 10

```

QY      4 PVPQLQPNPSQGGQPEQVPLVQ--QQQPFQGG--QGFPPQGPYPQPPSPQPYLQGL 59
DB      495 PLQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGG 553
QY      60 P---PFGQ---PFLPYPQ---PQSFPPQGPYPQ--QGYSGPQGPISQQAQGGQGGQGG 109
DB      554 PGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGG 613
QY      110 QGGQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGL 169
DB      614 EQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGG 663
QY      170 QALHNVAHAIILHQGGKQGGQSSQVSGFQGLQGYPLGGGFRPSPQGNPQAQGSVPQGL 229
DB      664 EQGGD-----EQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGG 716
QY      230 PQFEE 234
DB      717 EQGGD 721

```

RESULT 14

AAB62331 AAB62331 standard; protein; 1162 AA.

AAB62331;

06-AUG-2003 (revised)
29-JUN-2001 (first entry)

Amino acid sequence of KSHV tethering protein LANA.

Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KSHV; latency-associated nuclear antigen; LANA.

Human herpesvirus 8.

WO200125484-A2.

12-APR-2001.

29-SEP-2000; 2000WO-US026908.

01-OCT-1999; 99US-00410399.

(UNMI) UNIV MICHIGAN.

Robertson BS, Cotter MA;

WPI; 2001-281736/29.

N-PDB; AAF82901.

A composition for use in gene therapy comprises an expression vector that includes a nucleic acid sequence encoding a nucleic acid binding protein.

Disclosure; Fig 9B; 60pp; English.

The invention provides a composition comprising nucleic acid, histone H1 protein and expression vector operatively encoding a protein suitable for tethering the nucleic acid to the histone H1 protein, where the tethering protein is LANA. The composition is useful in aiding the retention of the viral DNA in the host cell. The viral vector encodes a protein suitable for tethering DNA to Histone H1. Methods for screening for compounds which are agonistic or antagonistic for the tethering of viral proteins to histone H1 and DNA binding sites are useful for developing the method of viral transfer. The composition has applications to gene therapy, including the treatment of multiple sclerosis, Parkinson's disease, Huntington disease and diabetes. The present sequence represents the amino acid sequence of the Kaposi's sarcoma associated herpesvirus (human herpesvirus 8) latency-associated nuclear antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-

CC 2003 to correct OS field.)

XX Sequence 1162 AA;

Query Match 18.6%; Score 266; DB 4; Length 1162;
Best Local Similarity 38.0%; Pred. No. 3, 8e-14;
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY      4 PVPQLQPNPSQGGQPEQVPLVQ--QQQPFQGG--QGFPPQGPYPQPPSPQPYLQGL 59
DB      495 PLQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGGQEPQGG 553
QY      60 P---PFGQ---PFLPYPQ---PQSFPPQGPYPQ--QGYSGPQGPISQQAQGGQGGQGG 109
DB      554 PGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGGREPPQGG 613
QY      110 QGGQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGLLQGL 169
DB      614 EQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGG 663
QY      170 QALHNVAHAIILHQGGKQGGQSSQVSGFQGLQGYPLGGGFRPSPQGNPQAQGSVPQGL 229
DB      664 EQGGD-----EQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGGDEGGQGG 716
QY      230 PQFEE 234
DB      717 EQGGD 721

```

RESULT 15

ABB05621 ABB05621 standard; protein; 1162 AA.

ABB05621;

25-APR-2002 (first entry)

Kaposi's sarcoma-associated herpesvirus LANA protein.

Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;
KSHV terminal repeat; rhadino virus cis acting element; episome;
primary effusion lymphoma; latency-associated nuclear antigen;
gene therapy; gene transfer.

Human herpesvirus 8.

US6322792-B1.

27-NOV-2001.

21-APR-1999; 99US-00298568.

19-NOV-1998; 98US-0109422P.

(KIEF/) KIEFF E D.

Kieff ED, Ballesstas ME, Kaye KM;

WPI; 2002-153769/20.

N-PDB; ABA93487.

System for episomal retention of plasmids in mammalian cells, useful in gene therapy, comprises rhadinoviral LANA and RVCAE sequences.

Disclosure; Fig 7; 27pp; English.

The present invention describes a system (A) for maintaining a plasmid as an episome in mammalian cells, comprising the rhadinoviral sequence LANA (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487, CC 51) expressed in the cell, and the rhadinoviral sequence RVCAE (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, 52) present in the plasmid. Also described is a method for maintaining a closed circular DNA in a cell by expressing (51) in the cells and having

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 seconds

(without alignments)
928.452 Million cell updates/sec

Title: US-10-089-700-3-Y65

Perfect score: 1433

Sequence: 1 VRVVPQLQPNPQQQPFQ.....CNVYAPYCTTAPGIFGTN 266

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	465	32.4	369	2	US-08-991-300-2
2	266	18.6	1162	2	US-08-728-323A-2
3	266	18.6	1162	3	US-09-298-568-2
4	266	18.6	1162	4	US-09-410-399-2
5	266	18.6	1162	4	US-09-894-273-2
6	266	18.6	1162	4	US-08-918-914-4
7	238	16.6	498	4	US-09-270-767-45042
8	234	16.3	256	4	US-09-248-796A-21251
9	219	15.3	2074	4	US-09-491-356C-9
10	218	15.2	579	4	US-09-668-119-3
11	201	14.0	2023	4	US-09-491-356C-8
12	201	14.0	2124	4	US-09-538-092-1377
13	199	13.9	663	4	US-09-270-767-61220
14	199	13.9	1591	4	US-09-270-767-45698
15	198.5	13.9	505	4	US-09-248-796A-19253
16	197	13.7	379	4	US-09-248-796A-19259
17	196.5	13.7	2441	1	US-08-194-468-2
18	196.5	13.7	2441	3	US-08-961-739-2
19	196.5	13.7	2441	3	US-09-514-247A-8
20	196.5	13.7	2441	4	US-09-514-247A-8
21	196	13.7	216	4	US-09-248-796A-21017
22	195.5	13.6	2442	3	US-09-514-247A-10
23	195.5	13.6	2442	4	US-09-538-092-1370
24	194.5	13.6	729	4	US-09-625-188-20
25	186	13.0	295	4	US-09-248-796A-20004
26	185.5	12.9	330	4	US-09-248-796A-24758
27	184.5	12.9	332	4	US-09-248-796A-21649

28	183.5	12.8	316	4	US-09-270-767-42663	Sequence 42663, A
29	183	12.8	169	4	US-09-248-796A-28087	Sequence 28087, A
30	183	12.8	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	848	4	US-09-538-092-33	Sequence 33, Appl
34	177	12.4	261	4	US-09-602-565-34	Sequence 34, Appl
35	177	12.4	1319	4	US-09-538-092-1281	Sequence 1291, Ap
36	176	12.3	618	4	US-09-248-796A-15319	Sequence 15319, A
37	175.5	12.2	684	4	US-09-823-240A-9	Sequence 9, Appl
38	175	12.2	657	4	US-09-248-796A-19232	Sequence 19232, A
39	174.5	12.2	382	4	US-09-248-796A-18720	Sequence 18720, A
40	174	12.1	1507	4	US-09-914-259-37	Sequence 37, Appl
41	173.5	12.1	675	4	US-09-248-796A-20699	Sequence 20699, A
42	172	12.0	542	1	US-07-814-964-13	Sequence 13, Appl
43	172	12.0	542	1	US-08-258-442-13	Sequence 13, Appl
44	172	12.0	542	1	US-08-328-809-8	Sequence 8, Appl
45	172	12.0	542	4	US-08-866-840-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-991-300-2
Sequence 2, Application US/08991300
Patent No. 5973225
GENERAL INFORMATION:
APPLICANT: D'OVIDIO, RENATO
APPLICANT: PORCEDDU, ENRICO
APPLICANT: MERCHITELLI, CINZIA
APPLICANT: CARRELLI, LUISA ERCOLI
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE
ENCODING A LOW MOLECULAR WEIGHT GLUTENIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,300
FILING DATE: 16-DEC-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT MI 96/A 002663
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2264-0201-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-300-2
Query Match 32.4%, Score 465, DB 2, Length 369;
Best Local Similarity 40.2%, Pred. No. 1.6e-35;

Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;

```

QY 6 PQLQPLQPNPSSQGP-----QGVPLVQGGQPGGQQGPP--PQGPVPPQPP----- 50
Db 38 PQQQPSCQQQQQPPVLSQQQQPPPSQQQQPPPSQQQQPPVLPQPSFQQQLPPSSQQQQPP 97
QY 51 ---SQEYL-----QLPFPQPLPYPPQSPFPQQ-----PYQPQPPQYSQ 89
Db 98 FSGQQQVPLPQQPFSQQQLPPFSQQQLPPFSQQQLPPFSQQQLPPFSQQQLPPFSQQ 157
QY 90 PQGPISQQQAQQQQQQ-----QQQQQQIQQ-----LQQQLPFCMVLVQ 132
Db 158 QQQQVPLPQQPFSQQQQQPPPPQQPFSQQQQQVPLVQQLPFPVPSILQQLNFC-KVFLQ 216
QY 133 QH-----NIAHARSQVLAQSTYQLLQELCCQHLWQIPESQSCQAIHNVVHAILHQK 186
Db 217 QGSPWAMPQSILARSQVLAQSSCHVQGCCQQLPQIFQSSRYEALRAIVYSILL--QEQ 274
QY 187 QQQQSSQVSPQQLQGYPLGQGSFRESQNPQAQGS-----VQPPQLPQFEE 234
Db 275 QQVQGSITQTDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQ 331
QY 235 IRLALQTLFAMCNVYIAPY--CTIAPFGI FGT 265
Db 332 MTSIALKTLPTMCNWNVPLVTRTVFVGT 363

```

RESULT 2

US-08-728-323A-2

Sequence 2, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan
 APPLICANT: Bohenzky, Roy A.
 APPLICANT: Russo, James J.
 APPLICANT: Edelman, Isidore S.
 APPLICANT: Moore, Patrick S.
 TITLE OF INVENTION: Immediate Early Protein From Kaposi's
 TITLE OF INVENTION: Sarcoma-Associated Herpesvirus' DNA
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/728,323A
 FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-278-0400
 TELEFAX: 212-391-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1162 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-728-323A-2

Query Match

18.6%; Score 266; DB 2; Length 1162;

Best Local Similarity 38.0%; Pred. No. 2,3e-16;
 Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY 4 PVPQLQPNPSSQGPQGEVPLVQ--QQQFPQQQ--QQFPQQPYPPQPPPSQQPYLQLQ 59
Db 495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEP 553
QY 60 P---FPQ---PYLPYPQ---PQSPFPQPPYPPQ-QPQYSQPPQPSISQQQAQQQQQQQQ 109
Db 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPP 613
QY 110 QQQQILQQILQQQLIFCMQDVVLQGHNTAHARSQVLAQSTYQLLQELCCQHLWQIPESQSC 169
Db 614 EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 663
QY 170 QAHNVVHAILLHQKQKQQQSSQVSPQQLQGYPLGQGSFRESQNPQAQGSVPQQL 229
Db 664 EQQDQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
QY 230 PQFEE 234
Db 717 EQQDQ 721

```

RESULT 3

US-09-298-568-2

Sequence 2, Application US/09298568

Patent No. 6322792

GENERAL INFORMATION:

APPLICANT: Kleff, Elliott D.
 APPLICANT: Ballestas, Mary E.
 APPLICANT: Kaye, Kenneth M.
 TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
 TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
 FILE REFERENCE: 16412-10001R
 CURRENT APPLICATION NUMBER: US/09/298,568
 CURRENT FILING DATE: 1999-04-21
 EARLIER APPLICATION NUMBER: US 60/109,422
 EARLIER FILING DATE: 1998-11-19
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 2

TYPE: PRT
 LENGTH: 1162
 ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-2

Query Match 18.6%; Score 266; DB 3; Length 1162;
 Best Local Similarity 38.0%; Pred. No. 2,3e-16;
 Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY 4 PVPQLQPNPSSQGPQGEVPLVQ--QQQFPQQQ--QQFPQQPYPPQPPPSQQPYLQLQ 59
Db 495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEP 553
QY 60 P---FPQ---PYLPYPQ---PQSPFPQPPYPPQ-QPQYSQPPQPSISQQQAQQQQQQQQ 109
Db 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPP 613
QY 110 QQQQILQQILQQQLIFCMQDVVLQGHNTAHARSQVLAQSTYQLLQELCCQHLWQIPESQSC 169
Db 614 EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 663
QY 170 QAHNVVHAILLHQKQKQQQSSQVSPQQLQGYPLGQGSFRESQNPQAQGSVPQQL 229
Db 664 EQQDQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716
QY 230 PQFEE 234
Db 717 EQQDQ 721

```

RESULT 4


```

US-09-410-399-2
: Sequence 2, Application US//09410399
: Patent No. 6482587
: GENERAL INFORMATION:
: APPLICANT: Robertson, Erle S.
: APPLICANT: Cotter, Murray A.
: TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
: TITLE OF INVENTION: to Genomic Host DNA
: FILE REFERENCE: UM-03778
: CURRENT APPLICATION NUMBER: US/09/410,399
: CURRENT FILING DATE: 1999-10-01
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1162
: TYPE: PRT
: ORGANISM: Kaposi's sarcoma-associated herpesvirus
: US-09-410-399-2

```

[illegible][illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```
Db 1807 QQQPVPQGGRLRQQ--LQSSQMLGSSVHQMTPSSSYGLQTSLSPPSLQGYISYSH 1864
Qy 52 ----QQ-----PYLQLPFPFPYLPYPP-----QSPFPQPPYPPQ-----PQYS 88
Db 1865 VGLQHTGTPADPTRLHQGPSSGVYHQAPTYGHGLTSTQRFSGHQLQCTPMWMTPLSA 1924
Qy 89 QPQQ-----PISQQAQQQQQQQQQQQQQQQQQQQLLQQLLQQLLPCMDVYLQCHNIAHA 139
Db 1925 QGVQAVRSTSLPEQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 1975
Qy 140 RSCVLAQSTYQLLQELCCQHLWQIPESQCCAIHNVVAIILHQQKQQQQPSSQVSPQ 199
Db 1976 RQQQQQQ-----QMLQQQQQQQQQ-----QQQQQQQQQQQQQQQQQQ 2013
Qy 200 PLQQLPLGGSPFPSSQNPQAGSVQPPQLPFPFETRLNA 239
Db 2014 PHQQ-----QQQAPFPQPPQSPQFQRCGLQCTQQQQQCTA 2049
```

```
RESULT 10
US-09-668-119-3
; Sequence 3, Application US/09668119
; Patent No. 6768003
; GENERAL INFORMATION:
; APPLICANT: Solomon, William B
; APPLICANT: Abraham, Shaji
; TITLE OF INVENTION: Transcriptional adaptor protein
; FILE REFERENCE: 011.00250
; CURRENT APPLICATION NUMBER: US/09/668,119
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-668-119-3
```

```
Query Match 15.2%; Score 218; DB 4; Length 579;
Best Local Similarity 34.5%; Pred. No. 2.9e-12;
Matches 91; Conservative 12; Mismatches 87; Indels 74; Gaps 11;
```

```
Qy 8 LQPNSSQQQPEQVPLVQQQFPQQQQFPFPQPPYPP--QPSSQGPYQLQFPFPY 65
Db 134 LQQQQQQQQFQQQQQALQQQQQQQQQQQQQQQQQQQQFQAQSSAMQQQFAVWQQQQQLQQQQQQQQQH 193
Qy 66 LPPYQSPFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 125
Db 194 LKRLHQN--QQQIQQQQQQL---QRIALQLQQQQQQQQQQQQQQQQQQ--QQALESQ--- 241
Qy 126 CNDVVLQCHNIAHARSQVLAQSTYQLLQELCCQHLWQIPESQCCAIHNVVAIILHQQ 185
Db 242 -----PPIQQPMMQPPQPPPSQALPQ-----QLQCHNHTQH---HQP 276
Qy 186 KQQQQSSQVSPFPQPPYLPYPP-----LQGSFRPQ-----QNPQAQ 221
Db 277 PGPQP--PVAQNPQPPQLPQSPQTPVQAQALPQMLYTPPLKTVAPNVVQPPVQ 334
Qy 222 GSVQPPQLPFPFETRLNALQTLPA 245
Db 335 PQVQQQQ-----TAVQTAA 349
```

```
RESULT 11
US-09-491-356C-8
; Sequence 8, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Gains, Edward I.
; APPLICANT: Delist, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF X013
```

```
FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491,356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 2023
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-491-356C-8
```

```
Query Match 14.0%; Score 201; DB 4; Length 2023;
Best Local Similarity 29.2%; Pred. No. 5.1e-10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11;
```

```
Qy 7 QLPQNPSSQQQPEQVPLVQQQFPQQQ--QSPFPQPPY----- 43
Db 1747 QQQPVPQGGRLRQQ--LQSSQMLGSSVHQMTPSSSYGLQTSQGYTPVSHVGLQHT 1804
Qy 44 ----PQPPFPSPQPPYQLQLPFPYLPY-----QPSPFPQPPY----- 81
Db 1805 GPAGTWPPPSYSSQPPYSTPSTPVLVDPTRHLQRPSSGVYHQAPTYGHGLTSTQFS 1864
Qy 82 ----QPQPPYSP-----PQPPISQQAQQQQQQQQQQQQQQQLLQ-- 117
Db 1865 HQTQLQTPMTSTPMNSAGVQAVSTALLPQQQQQQQQQQQQQQQQQQQQQQQQQQQY 1924
Qy 118 -ILQQQLIPCMDVVLQCHNIAHARSQVLAQSTYQLLQELCCQHLWQIPESQCCAIHNVV 176
Db 1925 HIRQQ-----QQQLRQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ-- 1962
Qy 177 HAILHQQKQQ-----QPSSQVSPQ--QPLQGPYLPGGSPFPSSQNPQAGSVQPPQ 227
Db 1963 ----HQQQQQQQQAPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 2015
```

```
RESULT 12
US-09-538-092-1377
; Sequence 1377, Application US/09538092
; Patent No. 6783314
; GENERAL INFORMATION:
; APPLICANT: Gloc, Joic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: ChrapatsegFormat Version 0.9
; SEQ ID NO 1377
; LENGTH: 2124
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number Q93074
US-09-538-092-1377
```

```
Query Match 14.0%; Score 201; DB 4; Length 2124;
Best Local Similarity 29.2%; Pred. No. 5.1e-10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11;
```

```
Qy 7 QLPQNPSSQQQPEQVPLVQQQFPQQQ--QSPFPQPPY----- 43
```


Wed Dec 15 10:02:04 2004

us-10-089-700-3-y65.rai

Page 7

Search completed: December 14, 2004, 17:29:01
Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:09:46 ; Search time 64.8333 Seconds
(without alignments)
1465.441 Million cell updates/sec

Title: US-10-089-700-3-y65

Perfect score: 1433

Sequence: 1 VRVVPQLQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1343.5	93.8	287	US-10-739-930-9777	Sequence 9777, Ap
2	1242.5	86.7	319	US-10-739-930-9619	Sequence 9619, Ap
3	1239	86.5	298	US-10-739-930-9770	Sequence 9770, Ap
4	588	41.0	327	US-10-739-930-9623	Sequence 9623, Ap
5	556.5	38.8	282	US-10-474-955-101	Sequence 101, App
6	545.5	38.1	298	US-10-739-930-9621	Sequence 9621, Ap
7	522	36.5	279	US-10-474-955-97	Sequence 97, Appl
8	522	36.4	279	US-10-474-955-99	Sequence 99, Appl
9	519	36.2	279	US-10-474-955-98	Sequence 98, Appl
10	519	36.2	279	US-10-474-955-100	Sequence 100, App
11	446.5	31.2	307	US-10-739-930-9778	Sequence 9778, Ap
12	438.5	30.6	304	US-10-739-930-9782	Sequence 9782, Ap
13	382	26.7	244	US-10-739-930-9769	Sequence 9769, Ap

14	277	19.3	541	US-10-425-115-200100	Sequence 200100,
15	274	19.1	283	US-10-425-115-200097	Sequence 200097,
16	266	18.6	1162	US-09-894-273-2	Sequence 2, Appl
17	266	18.6	1162	US-10-294-804-2	Sequence 2, Appl
18	264.5	18.5	1069	US-10-161-927-54	Sequence 54, Appl
19	255	17.8	323	US-10-425-114-44003	Sequence 44003, A
20	253	17.7	905	US-10-451-467A-64	Sequence 64, Appl
21	246.5	17.2	390	US-10-424-559-270450	Sequence 270450,
22	239.5	16.7	229	US-10-425-114-41056	Sequence 41056, A
23	239	16.7	358	US-10-104-047-3710	Sequence 3710, Ap
24	231.5	16.2	148	US-10-465-217-15	Sequence 15, Appl
25	229	16.0	1044	US-10-425-114-72709	Sequence 72709, A
26	227.5	15.9	362	US-10-425-115-261231	Sequence 261231,
27	227	15.8	1173	US-10-437-963-144743	Sequence 144743,
28	226.5	15.8	192	US-10-425-114-52411	Sequence 52411, A
29	226	15.8	738	US-10-451-467A-238	Sequence 238, App
30	222	15.5	1236	US-09-763-787-109	Sequence 109, App
31	220.5	15.4	1645	US-10-263-929-176	Sequence 176, App
32	220	15.4	4952	US-10-051-874-56	Sequence 56, Appl
33	220	15.4	5008	US-10-051-874-166	Sequence 166, Appl
34	220	15.4	5159	US-10-085-198-112	Sequence 112, App
35	220	15.4	5282	US-10-051-874-167	Sequence 167, App
36	220	15.4	5282	US-10-051-874-167	Sequence 167, App
37	218.5	15.2	351	US-10-264-049-2693	Sequence 2693, Ap
38	217.5	15.2	1024	US-10-479-546-12	Sequence 12, Appl
39	217.5	15.2	1153	US-10-479-546-1	Sequence 1, Appl
40	214.5	15.0	944	US-10-425-114-55600	Sequence 55600, A
41	214.5	15.0	1351	US-10-122A-75147	Sequence 75147, A
42	212.5	14.8	1359	US-10-437-963-190032	Sequence 190032,
43	212	14.8	1362	US-10-282-122A-75965	Sequence 75965, A
44	212	14.8	1362	US-09-815-242-14009	Sequence 14009, A
45	207.5	14.5	1025	US-10-437-963-113116	Sequence 113116,

ALIGNMENTS

RESULT 1
US-10-739-930-9777
; Sequence 9777, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9777
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_238.p
US-10-739-930-9777

Query Match 93.8%; Score 1343.5; DB 17; Length 287;
Best Local Similarity 95.1%; Pred. No. 8.1e-99;
Matches 254; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

QY	1	VRVVPQLQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap
DB	21	VRVVPQLQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap
QY	61	FPQQLPQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap
DB	81	FPQQLPQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap
QY	121	QQLPQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap
DB	141	QQLPQPNPSCQPPQCE.....CNVYIAPCTIAPGIRGTN 266	Sequence 9777, Ap


```

; APPLICANT: Drifhout, Jan W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludwig, Solid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
; FILE REFERENCE: 2799/71244-PCT-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; CURRENT FILING DATE: 2003-10-13
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 98
LENGTH: 279
TYPE: PRT
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Amino acid sequence of GAMMA-4
US-10-474-955-98

Query Match          36.4%; Score 522; DB 17; Length 279;
Best Local Similarity 46.2%; Pred. No. 1.5e-33;
Matches 132; Conservative 26; Mismatches 76; Indels 52; Gaps 14;

4 PVPQLQPNPSCQQQGEVPLVQQQGFPGQQQGF--PQQPYQPQPFPSQQPYLQLQPF 61
18 PVP--QPHQPSQCP-----QCTFPQQTTPHQPPQQQFPQCP--PQQQLQPPQCP 66
62 P-QPYLPPYPO--PQSFPPQPPYPPQ-----PQYSQPPQPISSQQAQ-----QQQQQQQQQQ 110
67 PQQPQQPYPQ-----QPPQFPQTQPPQLFPQSQPPQPPQPPQPPQPPQPPQPPQ 122
111 QQQILQQLLQQQLPQMDVVLQGN-----IAHARSQVLQSTYQLQELCCGHLMOIPQ 166
123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDCVMKQSCQQLAQILPQQ 182
167 SQCCAIHNVVAIILHQQKQKQQPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVQ 225
183 LQCAIHHTVHSITIMQEQEQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230
226 PQQLPQFEIRNALQTLTPAKCNVYIAPYCTI--APF-----GIFG 264
231 PQQPAQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

RESULT 9
US-10-474-955-98
Query Match          36.2%; Score 519; DB 17; Length 279;
Best Local Similarity 46.0%; Pred. No. 2.6e-33;
Matches 131; Conservative 26; Mismatches 77; Indels 52; Gaps 14;

4 PVPQLQPNPSCQQQGEVPLVQQQGFPGQQQGF--PQQPYQPQPFPSQQPYLQLQPF 61
18 PVP--QPHQPSQCP-----QCTFPQQTTPHQPPQQQFPQCP--PQQQLQPPQCP 66
62 P-QPYLPPYPO--PQSFPPQPPYPPQ-----PQYSQPPQPISSQQAQ-----QQQQQQQQQQ 110
67 PQQPQQPYPQ-----QPPQFPQTQPPQLFPQSQPPQPPQPPQPPQPPQPPQPPQ 122
111 QQQILQQLLQQQLPQMDVVLQGN-----IAHARSQVLQSTYQLQELCCGHLMOIPQ 166
123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDCVMKQSCQQLAQILPQQ 182
167 SQCCAIHNVVAIILHQQKQKQQPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVQ 225
183 LQCAIHHTVHSITIMQEQEQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230
226 PQQLPQFEIRNALQTLTPAKCNVYIAPYCTI--APF-----GIFG 264
231 PQQPAQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

FEATURE:
; OTHER INFORMATION: Amino acid sequence of GAMMA-2
US-10-474-955-98

Query Match          36.2%; Score 519; DB 17; Length 279;
Best Local Similarity 45.8%; Pred. No. 2.6e-33;
Matches 131; Conservative 26; Mismatches 77; Indels 52; Gaps 14;

4 PVPQLQPNPSCQQQGEVPLVQQQGFPGQQQGF--PQQPYQPQPFPSQQPYLQLQPF 61

```

```

DB 18 PVP--QPHQPSQCP-----QCTFPQQTTPHQPPQQQFPQCP--PQQQLQPPQCP 66
QY 62 P-QPYLPPYPO--PQSFPPQPPYPPQ-----PQYSQPPQPISSQQAQ-----QQQQQQQQQQ 110
DB 67 PQQPQQPYPQ-----QPPQFPQTQPPQLFPQSQPPQPPQPPQPPQPPQPPQ 122
QY 111 QQQILQQLLQQQLPQMDVVLQGN-----IAHARSQVLQSTYQLQELCCGHLMOIPQ 166
DB 123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDCVMKQSCQQLAQILPQQ 182
QY 167 SQCCAIHNVVAIILHQQKQKQQPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVQ 225
DB 183 LQCAIHHTVHSITIMQEQEQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230
QY 226 PQQLPQFEIRNALQTLTPAKCNVYIAPYCTI--APF-----GIFG 264
DB 231 PQQPAQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

RESULT 10
US-10-474-955-100
Query Match          36.2%; Score 519; DB 17; Length 279;
Best Local Similarity 46.0%; Pred. No. 2.6e-33;
Matches 131; Conservative 28; Mismatches 76; Indels 50; Gaps 14;

4 PVPQLQPNPSCQQQGEVPLVQQQGFPGQQQGF--PQQPYQPQPFPSQQPYLQLQPF 61
18 PVP--QPHQPSQCP-----QCTFPQQTTPHQPPQQQFPQCP--PQQQLQPPQCP 66
62 P-QPYLPPYPO--PQSFPPQPPYPPQ-----PQYSQPPQPISSQQAQ-----QQQQQQQQQQ 111
67 PQQPQQPYPQ-----QPPQFPQTQPPQLFPQSQPPQPPQPPQPPQPPQPPQPPQ 123
112 QQQILQQLLQQQLPQMDVVLQGN-----IAHARSQVLQSTYQLQELCCGHLMOIPQ 167
124 PPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDCVMKQSCQQLAQILPQQ 183
168 SQCCAIHNVVAIILHQQKQKQQPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVQ 226
184 LQCAIHHTVHSITIMQEQEQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 231
227 QQLPQFEIRNALQTLTPAKCNVYIAPYCTI--APF-----GIFG 264
232 QQPQQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

RESULT 11
US-10-739-930-9778
Query Match          36.2%; Score 519; DB 17; Length 279;
Best Local Similarity 46.0%; Pred. No. 2.6e-33;
Matches 131; Conservative 28; Mismatches 76; Indels 50; Gaps 14;

4 PVPQLQPNPSCQQQGEVPLVQQQGFPGQQQGF--PQQPYQPQPFPSQQPYLQLQPF 61
18 PVP--QPHQPSQCP-----QCTFPQQTTPHQPPQQQFPQCP--PQQQLQPPQCP 66
62 P-QPYLPPYPO--PQSFPPQPPYPPQ-----PQYSQPPQPISSQQAQ-----QQQQQQQQQQ 111
67 PQQPQQPYPQ-----QPPQFPQTQPPQLFPQSQPPQPPQPPQPPQPPQPPQPPQ 123
112 QQQILQQLLQQQLPQMDVVLQGN-----IAHARSQVLQSTYQLQELCCGHLMOIPQ 167
124 PPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDCVMKQSCQQLAQILPQQ 183
168 SQCCAIHNVVAIILHQQKQKQQPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVQ 226
184 LQCAIHHTVHSITIMQEQEQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 231
227 QQLPQFEIRNALQTLTPAKCNVYIAPYCTI--APF-----GIFG 264
232 QQPQQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

```

Query Match	30.6%	Score 438.5	DB 17	Length 307
Best Local Similarity	42.8%	Pred. No. 7,1e-27		
Matches	125	Conservative	36	Mismatches 76
				Indels 55
				Gaps 16
QY	13	PSQQQQLQEVLYQQQQLP-----QQQQQPPQQQLQVPPQSPQQPPIYLQLPQPPPIYL	67	
DB	27	PLGLERWQQQQQLPPLPQQTFFPQQLFSSQQQ--QQLFPPQPSSSQQP-----PFIQQQQLP	78	
QY	68	YVQPSFPQPQ-----YVQPSQSPQSPFI--SSQQAQQQQQQQQQQQQQQQI	114	
DB	79	FSQQQQLPILQQLPPFSQQQQLVLVSLQ--QPPSSQQQQLVLPQSSPFPQQQQQHQQLVQQQLP	137	

```

RESULT 14
US-10-425-115-200100
; Sequence: 200100, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21 (5322) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 369326
;

```


A:Reference number: A03354; MUID:84261434; PMID:6204862
A:Accession: A03354
A:Molecule type: DNA
A:Residues: 1-286 <RAP>
A:Cross-references: UNIPROT:P02863; GB:X00627; GB:X03076; NID:g21752; PIDN:CAA25261.1; F
A:Experimental source: cv. Newton
A:Comment: Gliadin is the major seed storage protein in wheat.
C:Superfamily: gliadin
C:Keywords: storage protein; tandem repeat
E:1-30/Domain: signal sequence #status predicted <SIG>
E:21-286/Product: gliadin #status predicted <GLN>
E:116-133/Region: glutamine-rich

Query Match 94.5%; Score 1354; DB 1; Length 286;
Best Local Similarity 95.5%; Pred. No. 1e-83;
Matches 254; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 60
DB 21 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 80
QY 61 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 120
DB 81 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 140
QY 121 QQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHAII 180
DB 141 QQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHAII 200
QY 181 LHQOQKQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQFEIRNLAL 240
DB 201 LHQOQKQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQFEIRNLAL 260
QY 241 QTLPMKCNVYIAPYCTIAPFGIFGTN 266
DB 261 QTLPMKCNVYIAPYCTIAPFGIFGTN 286

RESULT 3
106282
alpha-gliadin precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
A:Accession: T06282
R:Anderson, O.D.
Submitted to the EMBL Data Library, March 1996
A:Reference number: Z15587
A:Accession: T06282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-288 <AND>
A:Cross-references: UNIPROT:Q41530; EMBL:U51304; NID:g1256787; PIDN:AAA96523.1; PID:g125
A:Experimental source: cv. Cheyenne
C:Superfamily: gliadin
C:Keywords: seed; storage protein

Query Match 91.4%; Score 1310; DB 2; Length 288;
Best Local Similarity 92.5%; Pred. No. 9.1e-81;
Matches 248; Conservative 3; Mismatches 15; Indels 2; Gaps 1;

QY 1 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 60
DB 21 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 80
QY 61 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 118
DB 81 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 140
QY 119 LQOQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHA 178
DB 141 LQOQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHA 200

QY 179 IILHQQKQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQFEIRNL 238
DB 201 IILHQQKQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQFEIRNL 260
QY 239 ALQTLPMKCNVYIAPYCTIAPFGIFGTN 266
DB 261 ALQTLPMKCNVYIAPYCTIAPFGIFGTN 288

RESULT 4
T06498
alpha/beta-gliadin A-II precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
A:Accession: T06498
R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
A:Reference number: A92541; MUID:85234522; PMID:2989281
A:Accession: T06498
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <OKT>
A:Cross-references: UNIPROT:P04722; EMBL:M10092; NID:g170711; PIDN:AAA34276.1; PID:g17
C:Superfamily: gliadin
C:Keywords: seed; storage protein
E:1-20/Domain: signal sequence #status predicted <SIG>
E:21-291/Product: alpha/beta-gliadin A-II #status predicted <MAT>

Query Match 88.3%; Score 1265.5; DB 2; Length 291;
Best Local Similarity 87.6%; Pred. No. 8.6e-78;
Matches 240; Conservative 7; Mismatches 16; Indels 11; Gaps 2;

QY 1 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 60
DB 21 VRVVPVLQPNPESQOPQEVPLVQOQPFPGQOQPFPPQPPYPPQPPPSQPPYQLQLP 80
QY 61 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 120
DB 81 PPQPPYPPQPPPSQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQPPYPPQ 137
QY 121 QQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHAII 180
DB 138 QQLIPCMQVVLQGNHIAHRSQVLAQSTYQLQLQELCCQHLMOIPESQCCAIHNVVHAII 197
QY 181 LH-----QQKQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQ 232
DB 198 LHQOQHHAHQOQOQSSQSVSQPLQOQYPLQGSFPPSPQPPQAGSVQPPQLPQ 257
QY 233 EIRNLALQTLPMKCNVYIAPYCTIAPFGIFGTN 266
DB 258 EIRNLALQTLPMKCNVYIAPYCTIAPFGIFGTN 291

RESULT 5
S10015
alpha/beta-gliadin precursor (clone MW1) - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
A:Accession: S10015
R:Garcia-Maroto, F.; Marana, C.; Garcia-Olmedo, F.; Carbonero, F.
Plant Mol. Biol. 14, 867-868, 1990
A:Title: Nucleotide sequence of a cDNA encoding an alpha/beta-type gliadin from hexa
A:Reference number: S10015; MUID:91346679; PMID:2102865
A:Accession: S10015
A:Molecule type: mRNA
A:Residues: 1-307 <GAR>
A:Cross-references: UNIPROT:P18573; EMBL:X17361; NID:g21672; PIDN:CAA35238.1; PID:g216
C:Superfamily: gliadin
E:1-20/Domain: signal sequence #status predicted <SIG>
E:21-307/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 88.3%; Score 1265.5; DB 2; Length 307;

Best Local Similarity 84.7%; Pred. No. 9e-78; Indels 21; Gaps 3;
Matches 243; Conservative 8; Mismatches 15;

QY 1 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 60
DB 21 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 80
QY 61 F-----PQVPLPPOQSFPPQOQPPPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 105
DB 81 FPQOPLPYPQOPLPYPQOPLPYPQOPLPYPQOPLPYPQOPLPYPQOPLPYPQOPLPYPQ 140
QY 106 ---QQOQ 163
DB 141 QKQOQ 200
QY 164 PQSQOQ 219
DB 201 PQSQOQ 260
QY 220 AQSQVQOQ 266
DB 261 AQSQVQOQ 307

RESULT 6

alpha/beta-gliadin precursor (clone pm1215) - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 08-Jun-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S07361
R:Summer-Smith, M.; Rafalaki, J.A.; Sugiyama, T.; Stoll, M.; Soell, D.
Nucleic Acids Res. 13, 3905-3916, 1985
A:Title: Conservation and variability of wheat alpha/beta-gliadin genes.
A:Reference number: S07361; PMID:85242077; PMID:3839304
A:Accession: S07361
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <SUM>
A:Cross-references: UNIPROT:P04726; EMBL:X02538; NID:g21756; FIDN:CAA26383.1; PID:g21757
C:Superfamily: gliadin
C:Keywords: seed; storage protein

Query Match 88.0%; Score 1261; DB 2; Length 296;
Best Local Similarity 88.1%; Pred. No. 1.7e-77;
Matches 244; Conservative 4; Mismatches 17; Indels 12; Gaps 4;

QY 1 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 60
DB 21 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 80
QY 61 FPQO-----PQVPLPPOQSFPPQOQPPPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 111
DB 81 FPQOQ 140
QY 112 QOQLOQ 171
DB 141 QOQLOQ 200
QY 172 HNVVVAIILHQQOQ 231
DB 201 HNVVVAIILHQQOQ 260
QY 232 FEIRNIALQTLPRMCMNVYIAPYCTIAPFGIFGTN 266
DB 260 FEIRNIALQTLPRMCMNVYIAPYCTIAPFGIFGTN 296

RESULT 7

S07924
alpha/beta-gliadin precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 08-Jun-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S07924; C61218

R:Summer-Smith, M.; Rafalaki, J.A.; Sugiyama, T.; Stoll, M.; Soell, D.
Nucleic Acids Res. 13, 3905-3916, 1985
A:Title: Conservation and variability of wheat alpha/beta-gliadin genes.
A:Reference number: S07361; PMID:85242077; PMID:3839304
A:Accession: S07924
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-313 <SUM>

A:Cross-references: UNIPROT:Q41546; EMBL:X02540; NID:g21764; FIDN:CAA26385.1; PID:g2176
R:Shewry, P.R.; Sabelli, P.A.; Parmar, S.; Lafandra, D.
Biochem. Genet. 29, 207-211, 1991
A:Title: alpha-type prolamins are encoded by genes on chromosomes 4Ha and 6Ha of Hexyl
A:Reference number: A61218; PMID:91315394; PMID:1859356
A:Accession: C61218
A:Status: preliminary
A:Molecule type: protein
A:Residues: 18-27 <SHE>
C:Superfamily: gliadin
C:Keywords: seed; storage protein

Query Match 87.9%; Score 1259; DB 2; Length 313;
Best Local Similarity 82.8%; Pred. No. 2.5e-77;
Matches 245; Conservative 8; Mismatches 13; Indels 30; Gaps 4;

QY 1 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 60
DB 18 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 77
QY 61 FPQO-----PQVPLPPOQSFPPQOQPPPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 112
DB 78 FPQOQ 137
QY 113 QOQLOQ 172
DB 138 QOQLOQ 197
QY 173 HNVVVAIILH-----QQOQ 212
DB 198 HNVVVAIILHMQOQ 257
QY 213 PQSQOQ 266
DB 258 PQSQOQ 313

RESULT 8

A27319
gliadin - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 04-Mar-1988 #sequence_revision 04-Mar-1988 #text_change 03-Feb-1994
C:Accession: A27319
R:Reeves, C.D.; Ohta, T.W.
Gene 52, 257-266, 1987
A:Title: Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats.
A:Reference number: A27319; PMID:87277398; PMID:3038689
A:Accession: A27319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <REE>
C:Superfamily: gliadin

Query Match 87.6%; Score 1255; DB 2; Length 296;
Best Local Similarity 87.7%; Pred. No. 4.4e-77;
Matches 243; Conservative 3; Mismatches 19; Indels 12; Gaps 4;

QY 1 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 60
DB 21 VRVPVPOLOPQPNPSQOQPOEQVPLVQOQFPQOQOQFPPOQPYPOQPPFSQOQPYLOLP 80
QY 61 FPQO-----PQVPLPPOQSFPPQOQPPPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 111
DB 81 FPQOQ 140

D22364
alpha/beta-gliadin precursor (clone A735) - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C/Accession: D22364
R/Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A/Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
A/Reference number: A92541; MUID:85234522; PMID:2989281
A/Accession: D22364
A/Molecule type: mRNA
A/Residues: 1-326 <OKI>
A/Cross-references: UNIPROT:P04724
C/Superfamily: gliadin
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-326/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 85.2%; Score 1221; DB 2; Length 326;
Best Local Similarity 77.8%; Pred. No. 9e-75;
Matches 238; Conservative 9; Mismatches 19; Indels 40; Gaps 4;

QY 1 VRVVPQLQPNPNSQQQPOQEVPLV-QQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 60
DB 21 VRVVPQLQPNPNSQQQPOQEVPLVQQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 80
QY 61 F-----PPYLPYPPQPSFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 107
DB 81 FPPQPPFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 140
QY 108 QQQQQQILQQLIQQQLIPCMQVYLQGHNIHARSQVLTQSTYQLQLCCGHLMOIPQS 167
DB 141 QQQQQQILQQLIQQQLIPCRDVLVLOQHSIAHGSQVLTQSTYQLVQPCCCQQLWQIPQS 200
QY 168 QQQQHNVVHAIILH-----QQQKQQQPPSSQVSPQPLQ 202
DB 201 RQQAHHNVHAIILHQQQQQQQEQKQQLQQQQQQQQQQQQQQQQQQPPSLQVCPQSQ 260
QY 203 QYPLQGSFPPSQNPQAGSVQPPQLPQFEIRNLALQTLPMCNVTIAPYC--TIAPF 260
DB 261 QYPSQGSFPPSQNPQAGSVQPPQLPQFEIRNLALQTLPMCNVTIAPYCSTTIAPV 320
QY 261 GIFFGN 266
DB 321 GIFFGN 326

RESULT 13
E22364
alpha/beta-gliadin precursor (clone A1235) - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C/Accession: E22364
R/Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A/Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
A/Reference number: A92541; MUID:85234522; PMID:2989281
A/Accession: E22364
A/Molecule type: mRNA
A/Residues: 1-320 <OKI>
A/Cross-references: UNIPROT:P04723
C/Superfamily: gliadin
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-320/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 84.0%; Score 1204; DB 2; Length 320;
Best Local Similarity 79.0%; Pred. No. 1.2e-73;
Matches 237; Conservative 11; Mismatches 18; Indels 34; Gaps 5;

QY 1 VRVVPQLQPNPNSQQQPOQEVPLV-QQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 59
DB 21 VRVVPQLQPNPNSQQQPOQEVPLVQQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 80
QY 60 FPPQ-----PPLPYPPQPSFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 113

DB 81 FPPQPPFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 140
QY 114 ILQQLIQQQLIPCMQVYLQGHNIHARSQVLTQSTYQLQLCCGHLMOIPQS 173
DB 141 TLQQLIQQQLIPCRDVLVLOQHSIAHGSQVLTQSTYQLVQPCCCQQLWQIPQS 200
QY 174 NVVHAIILH-----QQQKQQQPPSSQVSPQPLQ 208
DB 201 NVVHAIILHHHQQQQEQKQQLQQQQQQQQQQQQQQQQQQPPSLQVSPQPPQ 260
QY 209 GSFPQGNPQAGSVQPPQLPQFEIRNLALQTLPMCNVTIAPYC--TIAPF 266
DB 261 VSFQSSQGNPQAGSVQPPQLPQFEIRNLALQTLPMCNVTIAPYCSTTIAPF 320

RESULT 14
T06504
alpha/beta-gliadin precursor (A-III) - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T06504
R/Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A/Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
A/Reference number: A92541; MUID:85234522; PMID:2989281
A/Accession: T06504
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-282 <OKI>
A/Cross-references: UNIPROT:P04723; EMBL:M11076; NID:G170725; PIDN:AAA34283.1; PID:G17C
C/Superfamily: gliadin
C/Keywords: seed; storage protein
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-282/Product: alpha/beta-gliadin A-III #status predicted <MAT>

Query Match 82.4%; Score 1181; DB 2; Length 282;
Best Local Similarity 84.8%; Pred. No. 3.7e-72;
Matches 228; Conservative 11; Mismatches 20; Indels 10; Gaps 4;

QY 1 VRVVPQLQPNPNSQQQPOQEVPLV-QQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 59
DB 21 VRVVPQLQPNPNSQQQPOQEVPLVQQQQPFQGGQQQFPFPQPPQPPSQQPYLQ 80
QY 60 FPPQPPFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 113
DB 81 FPPQPPFPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 133
QY 120 QQQQLIPCMQVYLQGHNIHARSQVLTQSTYQLQLCCGHLMOIPQSQCAIHNVTI 179
DB 134 QQQQLIPCRDVLVLOQHSIAHGSQVLTQSTYQLQLCCGHLMOIPQSRQAIHNVTI 193
QY 180 ILHQQQQQQPPSSQVSPQPLQYPLQGSFPPSQNPQAGSVQPPQLPQFEIRNL 239
DB 194 ILHHHQQQQQPPSSQVSPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 253
QY 240 LQTLPMCNVTIAPYC--TIAPF 266
DB 254 LQTLPMCNVTIAPYCSTTIAPF 320

RESULT 15
B22364
alpha/beta-gliadin precursor (clone A26) - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C/Accession: B22364
R/Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A/Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA
A/Reference number: A92541; MUID:85234522; PMID:2989281
A/Accession: B22364
A/Molecule type: mRNA

A:Residues: 1-292 <OKI>
A:Cross-references: UNIPROT:P04721
G:Superfamily: gliadin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-292/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 80.8%; Score 1157.5; DB 2; length 292;
Best Local Similarity 84.1%; Pred. No. 1,4e-70; Indels 31; Gaps 4;
Matches 227; Conservative 3; Mismatches 9;

```

  1 VRVVPQLQPNPSSQCPQCEQVPLVQQQCPFGQQQPPQQQYPPQPPSPQPPYLQLQP 60
  21 VRVVPQLQCNPEQQQPPQEQVPLVQQQCFLLGQQPPFPQQPYPPQPPSPQPPYLQLQP 80
  61 F--PPY---LPYPPQSPPPQPPYPPQPPQYSPQPPIS--QQQAQQQQQQQQQQQQQ 114
  81 FLQPPPPPPQLPYEQPPPPFPQPPQPPQYSPQPPISQQQQQQQQQQQQQQQQQQQ 140
  115 LQQILLQQQLPCMDVVLQOHNIHARSQVLQGSFYQLLQELCCQHLWQIPQSSQCAIHN 174
  141 IQQLLQQQLPCMDVVLQOHNIHARSQVLQGSFYQLLQELCCQHLWQIPQSSQCAIHN 200
  175 VVHAIIILH-----QQKQQQPSQSVSFQQPLQQYPLGGG 209
  201 VVHAIIILHQQQQQQEQKQLQQQQQQQQQLQQQQQQKQQQQPSQSVSFQQPLQQYPLGGG 260
  210 -SFRPSSQNPQKQSSVQPPQLPQFEIRNLA 239
  261 SFRPSSQNPQKQSSVQPPQLPQFEIRNLA 290
```

Search completed: December 14, 2004, 17:26:59
Job time : 15 secs

FT
VARIA

FT	VARIANT	193	194	HN -> LK (in Ref. 3)
SC	SEQUENCE	286 AA;	32949 MW;	ESECFABBB29E10C6 CRC64;
QY	Query Match	94.5%;	Score 1354;	DB 1; Length 286;
	Best Local Similarity	95.5%;	Pred. No. 4.1e-75;	
	Matches 254; Conservative	0;	Mismatches 12;	Indels 0; Gaps 0;
QY	1	VRVVPVLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	60	
Db	21	VRFPVPLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	80	
QY	61	FPQOQPYLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	120	
Db	81	FPQOQPYLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	140	
QY	121	QOLIPCMQVVLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHNVVAII	180	
Db	141	QOLIPCMQVVLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHNVVAII	200	
QY	181	LHQOQKQOQOQSSQVSVFQOPLQOYPLQOQSFPSQOQNPQOQSVQOQLOPQFEBIRNAL	240	
Db	201	LHQOQKQOQOQSSQVSVFQOPLQOYPLQOQSFPSQOQNPQOQSVQOQLOPQFEBIRNAL	260	
QY	241	QTLPAMCNVYIAPYCTIAPFGIGTN	266	
Db	261	QTLPAMCNVYIAPYCTIAPFGIGTN	286	
RESULT 3				
AAA96525	PRELIMINARY;	PRT;	286 AA.	
AC	AAA96525;			
DT	02-MAR-2004 (TREMBlrel. 27, Created)			
DT	02-MAR-2004 (TREMBlrel. 27, Last sequence update)			
DT	02-MAR-2004 (TREMBlrel. 27, Last annotation update)			
DE	Alpha-glucan storage protein.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.			
OX	NCBI_TaxId=4965;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=Cheyenne;			
RA	Anderson O.D.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U51307; AAA96525.1; ..			
SO	SEQUENCE 286 AA; 32949 MW; BSECBAABE29E10C6 CRC64;			
Query Match	94.5%;	Score 1354;	DB 2; Length 286;	
	Best Local Similarity	95.5%;	Pred. No. 4.1e-75;	
	Matches 254; Conservative	0;	Mismatches 12; Indels 0; Gaps 0;	
QY	1	VRVVPVLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	60	
Db	21	VRFPVPLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	80	
QY	61	FPQOQPYLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	120	
Db	81	FPQOQPYLQPNSSQOQPCQVPLVQOQCPFGQOQCPFPQOQPYQPOQPFPSQOQPYLQLOP	140	
QY	121	QOLIPCMQVVLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHNVVAII	180	
Db	141	QOLIPCMQVVLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHNVVAII	200	
QY	181	LHQOQKQOQOQSSQVSVFQOPLQOYPLQOQSFPSQOQNPQOQSVQOQLOPQFEBIRNAL	240	
Db	201	LHQOQKQOQOQSSQVSVFQOPLQOYPLQOQSFPSQOQNPQOQSVQOQLOPQFEBIRNAL	260	
QY	241	QTLPAMCNVYIAPYCTIAPFGIGTN	266	
Db	261	QTLPAMCNVYIAPYCTIAPFGIGTN	286	

FT	VARIANT	193	194	HN -> LK (in Ref. 3)
SC	SEQUENCE	286 AA;	32949 MW;	ESECFABBB29E10C6 CRC64;
QY	Query Match	94.5%;	Score 1354;	DB 1; Length 286;
	Best Local Similarity	95.5%;	Pred. No. 4.1e-75;	
	Matches 254; Conservative	0;	Mismatches 12;	Indels 0; Gaps 0;
QY	1	VRFVFPQLQONPSQQQPCFQVPLVQQQCPFGQQQCFPPQQPYPQDPQPFSSQDPYQLQCP	60	
Db	21	VRFVFPQLQONPSQQQPCFQVPLVQQQCPFGQQQCFPPQQPYPQDPQPFSSQDPYQLQCP	80	
QY	61	FPQPYLYFPQPSFPFPQPPYPQDPQPYQSPQCPQPSQQQQAQQQQQQQQQQQQQLQQQLQ	120	
Db	81	FPQPYLYFPQPSFPFPQPPYPQDPQPYQSPQCPQPSQQQQAQQQQQQQQQQQQQLQQQLQ	140	
QY	121	QQLIPCMQDVLLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHHNVAAII	180	
Db	141	QQLIPCMQDVLLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHHNVAAII	200	
QY	181	LHQQOKQKQPPSSQVSTFQPLQOYPLQGSFRRSQCNFPAQGSVQPOLQPFEBIRNAL	240	
Db	201	LHQQOKQKQPPSSQVSTFQPLQOYPLQGSFRRSQCNFPAQGSVQPOLQPFEBIRNAL	260	
QY	241	QTLPAMCNVYIAPYCTIAPFGIGTN	266	
Db	261	QTLPAMCNVYIAPYCTIAPFGIGTN	286	
RESULT 3				
AAA96525				
ID	AAA96525	PRELIMINARY;	PRT;	286 AA.
AC	AAA96525;			
DT	02-MAR-2004 (TREMBlrel. 27, Created)			
DT	02-MAR-2004 (TREMBlrel. 27, Last sequence update)			
DT	02-MAR-2004 (TREMBlrel. 27, Last annotation update)			
DE	Alpha-glucan storage protein.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.			
OX	NCBI_TaxId=4965;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=Cheyenne;			
RA	Anderson O.D.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U51307; AAA96525.1; .			
SR	SEQUENCE 286 AA; 32949 MW; B5ECBABB29E10C6 CRC64;			
QY	Query Match	94.5%;	Score 1354;	DB 2; Length 286;
	Best Local Similarity	95.5%;	Pred. No. 4.1e-75;	
	Matches 254; Conservative	0;	Mismatches 12;	Indels 0; Gaps 0;
QY	1	VRFVFPQLQONPSQQQPCFQVPLVQQQCPFGQQQCFPPQQPYPQDPQPFSSQDPYQLQCP	60	
Db	21	VRFVFPQLQONPSQQQPCFQVPLVQQQCPFGQQQCFPPQQPYPQDPQPFSSQDPYQLQCP	80	
QY	61	FPQPYLYFPQPSFPFPQPPYPQDPQPYQSPQCPQPSQQQQAQQQQQQQQQQQQQLQQQLQ	120	
Db	81	FPQPYLYFPQPSFPFPQPPYPQDPQPYQSPQCPQPSQQQQAQQQQQQQQQQQQQLQQQLQ	140	
QY	121	QQLIPCMQDVLLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHHNVAAII	180	
Db	141	QQLIPCMQDVLLQCHNIAHRSQVLAQSTYLLQELCCQHLMQIPESQCCAHHNVAAII	200	
QY	181	LHQQOKQKQPPSSQVSTFQPLQOYPLQGSFRRSQCNFPAQGSVQPOLQPFEBIRNAL	240	
Db	201	LHQQOKQKQPPSSQVSTFQPLQOYPLQGSFRRSQCNFPAQGSVQPOLQPFEBIRNAL	260	
QY	241	QTLPAMCNVYIAPYCTIAPFGIGTN	266	
Db	261	QTLPAMCNVYIAPYCTIAPFGIGTN	286	

RESULT 4

Q9ZP09 PRELIMINARY; PRT; 288 AA.
 ID Q9ZP09
 AC Q9ZP09
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-gliadin precursor (Fragment).
 GN Name=alpha-gliadin;
 OS Triticum aestivum subsp. spelta.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 CX NCBI_TaxID=58933;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kaaard D.D., D'Ovidio R.;
 RT "Deduced amino acid sequences of an alpha-gliadin gene from Spelt wheat
 (Spelta) includes sequences active in celiac disease.";
 RL Cereal Chem. 76:548-551(1999).
 DR EMBL; AJ130948; CAA10257.1; -.
 DR PIR; G13333; S13333.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; G1a_glutenin.
 DR Pfam; PF00234; TYP_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR PRINTS; PR00209; GLIADIN.
 DR SMART; SM00499; AAI; 1.
 DR SIGNAL.
 FT CHAIN 1 20 Potential.
 FT NON TER 288 alpha-gliadin.
 FT SEQUENCE 288 AA; 33203 MW; DA058F3FAFA6BC6C CRC64;
 SQ
 Query Match 94.3%; Score 1352; DB 2; Length 288;
 Best Local Similarity 95.1%; Pred. No. 5, 4e-75;
 Matches 255; Conservative 0; Mismatches 11; Indels 2; Gaps 1;
 QY 1 VRVPVQLQPONPSQQQPOEQVPLVQOQOFPQOQOFPPOQPYPOPPFPSSQOPLYLOP 60
 DB 21 VRVPVQLQPONPSQQQPOEQVPLVQOQOFPQOQOFPPOQPYPOPPFPSSQOPLYLOP 80
 QY 61 FPQPLPYPOPSFPPOQPYPOPPQYSPQPPIS--QQQAQQOQQOQQOQQOQQOQQO 118
 DB 81 FPQPLPYPOPSFPPOQPYPOPPQYSPQPPISQQOQQOQQOQQOQQOQQOQQOQQO 140
 QY 119 LQQQLIPCMQVYLQCHNIAHARSQVLAQSTYQLQLCCOHLMQIPESQCCOAIHNVHA 178
 DB 141 LQQQLIPCMQVYLQCHNIAHARSQVLAQSTYQLQLCCOHLMQIPESQCCOAIHNVHA 200
 QY 179 IILHQOQKQOQSSQVSPFQPLQOYPLGGSFRRSQNPQAQGSVQPOQLPQFEIINTL 238
 DB 201 IILHQOQKQOQSSQVSPFQPLQOYPLGGSFRRSQNPQAQGSVQPOQLPQFEIINTL 260
 QY 239 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 266
 DB 261 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 288
 PRELIMINARY; PRT; 274 AA.
 ID Q9M4M5
 AC Q9M4M5
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-gliadin
 GN Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.

RESULT 6

Q9M4M2 PRELIMINARY; PRT; 276 AA.
 ID Q9M4M2
 AC Q9M4M2
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-gliadin
 GN Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 CX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Arenz-Hansen E.H., McAdam S.N., Wolberg O., Kristiansen C.,
 RA Solild L.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ133606; CAB76958.1; -.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; G1a_glutenin.
 DR Pfam; PF00234; TYP_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR PRINTS; PR00209; GLIADIN.
 DR SMART; SM00499; AAI; 1.
 FT CHAIN 1 276 alpha-gliadin.
 SQ
 Query Match 94.1%; Score 1348.5; DB 2; Length 274;
 Best Local Similarity 94.1%; Pred. No. 8, 5e-75;
 Matches 255; Conservative 0; Mismatches 11; Indels 5; Gaps 1;
 QY 1 VRVPVQLQPONPSQQQPOEQVPLVQOQOFPQOQOFPPOQPYPOPPFPSSQOPLYLOP 60
 DB 2 VRVPVQLQPONPSQQQPOEQVPLVQOQOFPQOQOFPPOQPYPOPPFPSSQOPLYLOP 61
 QY 61 FPQPLPYPOPSFPPOQPYPOPPQYSPQPPIS----QQQAQQOQQOQQOQQOQQOQQO 115
 DB 62 FPQPLPYPOPSFPPOQPYPOPPQYSPQPPISQQOQQOQQOQQOQQOQQOQQOQQO 121
 QY 116 QQLLQOQLIPCMQVYLQCHNIAHARSQVLAQSTYQLQLCCOHLMQIPESQCCOAIHNV 175
 DB 122 QQLLQOQLIPCMQVYLQCHNIAHARSQVLAQSTYQLQLCCOHLMQIPESQCCOAIHNV 181
 QY 176 VHAIIHQOQKQOQSSQVSPFQPLQOYPLGGSFRRSQNPQAQGSVQPOQLPQFEIINTL 235
 DB 182 VHAIIHQOQKQOQSSQVSPFQPLQOYPLGGSFRRSQNPQAQGSVQPOQLPQFEIINTL 241
 QY 236 RNIALQTLPMQCNVYIAPYCTIAPFGIFGTN 266
 DB 242 RNIALQTLPMQCNVYIAPYCTIAPFGIFGTN 272

[illegible]

Query Match	93.2%	Score 1335.5	DB 2	Length 289
Best Local Similarity	94.1%	Pred. No. 5.5e-74		
Matches 253	Conservative	1	Mismatches 12	Indels 3
				Gaps 1
QY	1	VRVVPVLOLPQNPSSQCCQPCQVPLVYQQCCQPPGQQCCQFPFPQPYPPQPPFPSSQCPYQLQLP	60	
DB	21	VRVSVPLQQLQNPSSQCCQPCQVPLVYQQCCQFLGQQCCQFPFPQPYPLQLQFPSSQCPYQLQLP	80	
QY	61	FPQPYPLPQPPQSPFPQPPQPYPPQPPQYSQPPQPS---	117	
DB	81	FPQPYPLPQPPQSPFPQPPQPYPPQPPQYSQPPQPS---	140	
QY	118	ILQQQLPQMDVVLQGNHIAHARQVQSGSTYCLJELCCQHLMQIPESQCAIHNHVV	177	
DB	141	ILQQQLPQMDVVLQGNHIAHARQVQSGSTYCLJELCCQHLMQIPESQCAIHNHVV	200	
QY	178	ALIHQQCKQQQPPSSQVSPFQQLQYPLGGGSRFPSCQNPQAGSQVQPPQLPQFEIRN	237	
DB	201	ALIHQQCKQQQPPSSQVSPFQQLQYPLGGGSRFPSCQNPQAGSQVQPPQLPQFEIRN	260	
QY	238	LALQTLPMCNVYIAPYCTIAPFGIFGTN	266	
DB	261	LALQTLPMCNVYIAPYCTIAPFGIFGTN	289	
RESULT 11				
QSM4L9				
ID	QSM4L9	PRELIMINARY	PRT	270 AA.
AC	QSM4L9	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Alpha-glucanin.			
OS	Triticum aestivum (wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;			
OC	Triticaceae; Triticum.			
NCBI	NCB1	taxid=4565;		
RN	1	SEQUENCE FROM N.A.		
RP	STRAIN=McJelmer; TISSUE=Endosperm;			
RA	Aarettz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,			
RL	Sollid L.M.			
RL	Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; GI13609; CAB7696.1			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR0013612; AAI.			
DR	InterPro; IPR0013767; Glucanin.			
DR	InterPro; IPR001954; Glucanin.			
DR	Pfam; PF00234; TYP alpha-amyl.			
DR	PRINTS; PR00208; GLIADGLUTEN.			
DR	PRINTS; PR00209; GLINDIN.			
DR	SMART; SM00499; AAI; 1.			
FT	CHAIN	1	270	alpha-glucanin.
SEQ	SEQUENCE	270 AA;	31491 MW;	IDB4B528BFAFFS CRC64;
Query Match	92.7%	Score 1328.5	DB 2	Length 270
Best Local Similarity	94.0%	Pred. No. 1.4e-73		
Matches 251	Conservative	3	Mismatches 12	Indels 1
QY	1	VRVVPVLOLPQNPSSQCCQPCQVPLVYQQCCQPPGQQCCQFPFPQPYPPQPPFPSSQCPYQLQLP	60	
DB	2	VRVSVPLQQLQNPSSQCCQPCQVPLVYQQCCQFLGQQCCQFPFPQPYPLQLQFPSSQCPYQLQLP	61	
QY	61	FPQPYPLPQPPQSPFPQPPQPYPPQPPQYSQPPQPS---	119	
DB	62	FPQPYPLPQPPQSPFPQPPQPYPPQPPQYSQPPQPS---	121	

QY 120 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 179
 DB 122 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 181
 QY 180 ILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPOFEIRNLAL 239
 DB 182 ILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPOFEIRNLAL 241
 QY 240 LQTLPMQCNVYIAPYCTIAPFGIFGTN 266
 DB 242 LQTLPMQCNVYIAPYCTIAPFGIFGTN 268

RESULT 12

QY 120 Q9M4M1 PRELIMINARY; PRT; 278 AA.
 AC Q9M4M1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DB 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-glucan.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 FT CHAIN 1 278 alpha-glucan.
 SQ SEQUENCE 278 AA; 32502 MW; COEBSBFDIDCA87D CRC64;

Query Match 92.6%; Score 1327.5; DB 2; Length 278;
 Best Local Similarity 92.0%; Pred. No. 1.6e-73;
 Matches 253; Conservative 1; Mismatches 12; Indels 9; Gaps 1;
 QY 1 VRVVPQQLQPNPSCQQLQPEQVPLVQCCQFEGCCQCFPPQQLPQPPPSQQLQQLP 60
 DB 2 VRVVPQQLQPNPSCQQLQPEQVPLVQCCQFEGCCQCFPPQQLPQPPPSQQLQQLP 61
 QY 61 PPQPLPYPQPSPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 111
 DB 62 PPQPLPYPQPSPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 121
 QY 112 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 171
 DB 122 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 181
 QY 172 ILHNVHAIILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPO 231
 DB 182 ILHNVHAIILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPO 241
 QY 232 FEIRNLALQTLPMQCNVYIAPYCTIAPFGIFGTN 266
 DB 242 FEIRNLALQTLPMQCNVYIAPYCTIAPFGIFGTN 276

RESULT 13

QY 120 Q41533 PRELIMINARY; PRT; 259 AA.
 AC Q41533;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DB 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-glucan.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 FT CHAIN 1 259 alpha-glucan mature peptide.
 SQ SEQUENCE 259 AA; 29996 MW; F36CD48FD8F54C6 CRC64;

DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-glucan (Fragment).
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98409296; PubMed=9738916;
 RA Maruyama N., Ichise K., Katsube T., Kishimoto T., Kawase S.,
 RA Matsumura Y., Takeuchi Y., Sawada T., Utsuni S.,
 RT "Identification of major wheat allergens by means of the Escherichia
 coli expression system.";
 RL Eur. J. Biochem. 255:739-745(1998).
 DR GO:GO:0045735; F:nutrient reservoir activity; IEA.
 DR EMBL: D84341; BAA12318.1; -
 DR GO:GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001376; Gliadin.
 DR InterPro: IPR001376; Gliadin.
 DR Pfam: PF00234; Tryp_alpha.amyl.; 1.
 DR PRINTS: PR00208; GLIADGLUTEN.
 DR PRINTS: PR00209; GLIADIN.
 DR SMART: SM00499; AAI; 1.
 FT NON TER 1 259
 FT CHAIN 1 259
 SQ SEQUENCE 259 AA; 29996 MW; F36CD48FD8F54C6 CRC64;

Query Match 91.7%; Score 1313.5; DB 2; Length 259;
 Best Local Similarity 93.2%; Pred. No. 1.1e-72;
 Matches 248; Conservative 1; Mismatches 10; Indels 7; Gaps 1;
 QY 1 VRVVPQQLQPNPSCQQLQPEQVPLVQCCQFEGCCQCFPPQQLPQPPPSQQLQQLP 60
 DB 1 VRVVPQQLQPNPSCQQLQPEQVPLVQCCQFEGCCQCFPPQQLPQPPPSQQLQQLP 60
 QY 61 PPQPLPYPQPSPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 120
 DB 61 PPQPLPYPQPSPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 113
 QY 121 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 180
 DB 114 QOOLIPCMQVVLQOHNIHARSQVLTQSTYQLQELCCQHLWQIPESGCCAIHNVHAI 173
 QY 181 ILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPOFEIRNLAL 240
 DB 174 ILHQQKQKQKQSSQVSPQQLQGYPLGGGSRPSQONPQAGSVQPOLPOFEIRNLAL 233
 QY 241 QTLPMQCNVYIAPYCTIAPFGIFGTN 266
 DB 234 QTLPMQCNVYIAPYCTIAPFGIFGTN 259

RESULT 14

QY 120 Q41530 PRELIMINARY; PRT; 288 AA.
 AC Q41530;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Alpha-glucan storage protein.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Cheyenne;
 RA Anderson O.D.;
 RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U51304; AAA96523.1; -.
 DR PIR; S13333; S13333.
 DR PIR; T06282; T06282.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003512; AAT.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; Gliadin.
 DR Pfam; PF00234; TRY_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR PRINTS; PR00209; GLIADIN.
 DR SMART; SM00499; AAT; 1.
 SQ SEQUENCE 288 AA; 33285 MW; D7F6B99133283CA2 CRC64;

Query Match 91.4%; Score 1310; DB 2; Length 288;
 Best Local Similarity 92.5%; Pred. No. 2e-72;
 Matches 248; Conservative 3; Mismatches 15; Indels 2; Gaps 1;

QY 1 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLFGLQQQFPFPQCPYPQCPFPSPQCPYLQLP 60
 DB 21 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLFGLQQQFPFPQCPYPQCPFPSPQCPYLQLP 80
 QY 61 FPQPLPYSPQPFPRPQCPYPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPY 118
 DB 81 FPQPLPYSPQPFPRPQCPYPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPY 140
 QY 119 LQQQLPCMDVVLQGHNTAHASQVLAQSTYQLLELCCGHLWQIEQSQCAIHNVAHAI 178
 DB 141 LQQQLPCMDVVLQGHNTAHASQVLAQSTYQLLELCCGHLWQIEQSQCAIHNVAHAI 200
 QY 179 ILHQQK-QQQQPSQVSFQCPPLQYPLQGSFRPSQNPQAGSVQCPQLPQFEIRNL 238
 DB 201 ILHQQK-QQQQPSQVSFQCPPLQYPLQGSFRPSQNPQAGSVQCPQLPQFEIRNL 260
 QY 239 ALQTLPMCNVYIAPYCTIAPFGIFGTN 266
 DB 261 ALQTLPMCNVYIAPYCTIAPFGIFGTN 288

RESULT 15

Q41528 PRELIMINARY; PRT; 287 AA.

AC Q41528;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Alpha-gliadin.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cheyenne;
 RA Anderson O.D.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U50984; AAA96276.1; -.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR003512; AAT.
 DR InterPro; IPR001376; Gliadin.
 DR InterPro; IPR001954; Gliadin.
 DR Pfam; PF00234; TRY_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR PRINTS; PR00209; GLIADIN.
 DR SMART; SM00499; AAT; 1.
 SQ SEQUENCE 287 AA; 33193 MW; 05F82296749CE97 CRC64;

Query Match 90.5%; Score 1297.5; DB 2; Length 287;
 Best Local Similarity 92.9%; Pred. No. 1.1e-71;
 Matches 249; Conservative 2; Mismatches 14; Indels 3; Gaps 3;

QY 1 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLFGLQQQFPFPQCPYPQCPFPSPQCPYLQLP 60

DB 21 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLFGLQQQFPFPQCPYPQCPFPSPQCPYLQLP 79
 QY 61 FPQPLPYSPQPFPRPQCPYPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPY 119
 DB 80 FPQPLPYSPQPFPRPQCPYPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPYSPQCPY 139
 QY 120 LQQQLPCMDVVLQGHNTAHASQVLAQSTYQLLELCCGHLWQIEQSQCAIHNVAHAI 179
 DB 140 LQQQLPCMDVVLQGHNTAHASQVLAQSTYQLLELCCGHLWQIEQSQCAIHNVAHAI 199
 QY 180 ILHQQK-QQQQPSQVSFQCPPLQYPLQGSFRPSQNPQAGSVQCPQLPQFEIRNL 238
 DB 200 ILHQQK-QQQQPSQVSFQCPPLQYPLQGSFRPSQNPQAGSVQCPQLPQFEIRNL 259
 QY 239 ALQTLPMCNVYIAPYCTIAPFGIFGTN 266
 DB 260 ALQTLPMCNVYIAPYCTIAPFGIFGTN 287

Search completed: December 14, 2004, 17:25:22
 Job time : 80 secs

This Page Blank (uspto)